BULLETIN

OF THE

AMERICAN MATHEMATICAL SOCIETY

EDITED BY

EARLE R. HEDRICK
DAVID R. CURTISS WILLIAM R. LONGLEY

WITH THE ASSISTANCE OF

HAROLD T. DAVIS HARRY W. KUHN
THEOPHIL H. HILDEBRANDT DERRICK N. LEHMER
JOHN R. KLINE HENRY L. RIETZ

VOLUME XLI, NUMBER 12
DECEMBER, 1935

PUBLISHED BY THE SOCIETY
MENASHA, WIS., AND NEW YORK
1935

Whole No. 435 $9.00 a Year

PUBLISHED MONTHLY

Entered as second class matter April 24, 1926 at the post office at Menasha, Wis., under the act of August 24, 1912. Acceptance for mailing at the special rate of postage provided for in the act of February 28, 1925, embodied in paragraph 4, section 538, P. L. and R., authorized May 9, 1935.

PERIODICAL PUBLICATIONS
OF THE
AMERICAN MATHEMATICAL SOCIETY

BULLETIN OF THE
AMERICAN MATHEMATICAL SOCIETY

This journal is published monthly. It is the official organ of the Society. One of its functions is to report official acts of the Society and the details of its meetings, with abstracts of all research papers presented. In addition to these official reports, it publishes also short original papers on mathematical topics, and some of the longer officially invited addresses presented before the Society. These papers form the body of the Bulletin. It contains also announcements and notes on matters of current interest, and reviews of advanced mathematical books.

Price per annual volume, $9.00. Sets of early volumes are available as follows: vols. 12–36 (1905–1930), $7.00 each; vols. 37 to date, $9.00 each. It is proposed to reprint vols. 1–11 and the three volumes of the Bulletin of the New York Mathematical Society (the precursor of the present Bulletin), which are out of print and scarce; complete sets can be obtained through the Society's New York office as soon as this work has been finished.

TRANSACTIONS OF THE
AMERICAN MATHEMATICAL SOCIETY

This Journal is devoted to research in pure and applied mathematics.

Beginning with 1935, two volumes of three numbers each of the Transactions will be published per year. The price per volume will be $6.00, so that the annual subscription price (six issues) will be $12.00. The Transactions has appeared as a quarterly since 1900, with the exception of the year 1922, when two volumes, of four numbers each were published (vols. 23–24).

Sets of early volumes are available as follows: vols. 1–21 (1900–1920), $7.00 each; vols. 22–32 (1921–1930), $8.00 each; vols. 33–36 (1931–1934), $9.00 each.

Subscriptions and orders for the Bulletin or Transactions should be addressed to the American Mathematical Society, Menasha, Wisconsin, or Low Library, Columbia University, New York City.

ON THE COMPLEX ROOTS OF ALGEBRAIC EQUATIONS*

BY A. J. KEMPNER

1. *Introduction.* The topic I have chosen is so vast that a systematic treatment in an address is impossible. I shall, therefore, restrict my report to some phases of the problem to which I have made contributions, and to such papers as deal with related questions and which have come to my notice.

I sincerely appreciate this opportunity to bring together the results of several articles, all dealing more or less directly with the problem of complex roots, but scattered in various journals over a period of many years.

Our equations are always polynomials, equated to zero, and, unless otherwise stated, with real coefficients. The notation $f_n(z)=0$ indicates that the equation is of degree n. Merely to answer the question whether equations of high degree are ever actually solved, it may be pointed out that a table of natural sines and cosines, say for every $1''$, is nothing but a complete tabulation of the solutions of the equation $x^{1296000}-1=0$.

2. *On Equations with Roots $e^{i\theta}$.* We have in the literature many theorems on equations with roots of absolute value unity. One of the most interesting of these is Kronecker's theorem:

If the coefficients of an equation are ordinary integers, if the coefficient of the highest power is unity, and if all roots are of absolute value unity, then all the roots are roots of unity, and the equation is therefore solvable by radicals.

A companion theorem is as follows.

With the same restrictions on the coefficients, if all roots are real and of absolute value <2, the roots are all of the form $2\cos \pi k$, where k is rational.

In two short notes† I have derived necessary and sufficient conditions that an equation have some, or all, of its roots of the form $e^{i\theta}$.

---

* An address presented to the Society, by invitation of the program committee, Lincoln, Nebraska, November 30, 1934.

† Kempner, Archiv der Mathematik und Physik, (3), vol. 25 (1916), pp. 236–242. Kempner, Tôhoku Mathematical Journal, vol. 10 (1916), pp. 115–117.

THEOREM 1. *To test any equation $f_n(z)=0$ with real coefficients for roots of the form $e^{i\theta}$, form the following polynomial, which has real coefficients and which contains only even powers of z:*

$$\phi_{2n}(z) = (z^2+1)^n \cdot f\left(\frac{z+i}{z-i}\right) \cdot f\left(\frac{z-i}{z+i}\right) = \psi_n(z^2) = \psi_n(u).$$

The equation $f_n(z)=0$ has among its roots some of the form $e^{i\theta}$ when and only when $\psi_n(u)=0$ has some positive roots. A necessary and sufficient condition that all roots of $f_n(z)=0$ be of the form $e^{i\theta}$ is that $\psi_n(u)=0$ have only positive roots.

For proof we have only to remember that $z=(z_1+i)/(z_1-i)$ transforms the unit circle in the z plane into the axis of reals in the z_1 plane. The theorem may be extended in an obvious fashion to equations with complex coefficients, by considering $f(z)\cdot\overline{f(z)}=F(z)$ in place of $f(z)$. It is clear that the theorem applies to much wider classes of functions than polynomials. If $\psi_n(u)=0$ has only real positive roots, and if, in addition, the coefficients of $f_n(z)=0$ are real integers and the coefficient of the highest power is unity, the roots of $f_n(z)=0$ are roots of unity; and, by Kronecker's theorem, the equation $f_n(z)=0$ is solvable by radicals.

Making use of the inverse transformation $z_1=i(z+1)/(z-1)$ and retracing our steps, we now start from a given equation $\phi_n(z)=a_0z^n+a_1z^{n-1}+\cdots+a_n=0$ whose roots are all real, and form the new polynomial

$$\psi_{2n}(z) = (z-1)^{2n}\phi_n\left(i\,\frac{z+1}{z-1}\right)\phi_n\left(-\,i\,\frac{z+1}{z-1}\right).$$

If $\phi_n(z)$ has integral coefficients, the coefficients of $\psi_{2n}(z)$ are also integers. It is easily seen that the coefficient of z^{2n} in $\psi_{2n}(z)$ is $(a_0-a_2+a_4-+\cdots)^2+(a_1-a_3+a_5-+\cdots)^2=\gamma$, say. If $\gamma=1$, $\psi_{2n}(z)=0$ satisfies the conditions of Kronecker's theorem. Hence $\psi_{2n}(z)=0$ is solvable by radicals, and hence also $\phi_n(z)=0$ is solvable by radicals.

THEOREM 2. *Suppose that $\phi_n(z)=a_0z^n+a_1z^{n-1}+\cdots+a_n=0$ has real integral coefficients and that all its roots are real. If*

$(a_0-a_2+a_4-+\cdots)^2+(a_1-a_3+a_5-+\cdots)^2=1$, *the equation is solvable by radicals.**

3. *On Equations with Roots* $\alpha+\rho e^{i\theta}$, *where* α *and* ρ *are Rational.* Of greater interest is the investigation of irreducible equations admitting roots of the form $\alpha+\rho e^{i\theta}$, where α and ρ are both rational.† Hand in hand with these equations go those admitting roots of the form $\alpha+\rho e^{i\theta}$, where α and ρ^2 are rational.

We first characterize geometrically the set of points in the plane of complex numbers corresponding to the set of numbers $\alpha+\rho e^{i\theta}$, where α and ρ are rational. The numbers are represented by the doubly infinite set of circles whose centers are at rational points on the axis of reals, and whose radii are rational. We shall call this set of circles, or the corresponding set of numbers, the set S.

Similarly, we shall represent the set of numbers $\alpha+\rho e^{i\theta}$, where α and ρ^2 are rational, by a corresponding doubly infinite set of circles, and we shall call this set of circles, or the set of numbers, the set C.

We shall consider the class of irreducible equations that have *any* complex roots on any circle S. For simplicity, unless otherwise stated, we shall assume that the coefficients are integral and that the equation is irreducible in the natural domain. This class consists of:

(1) all irreducible equations $f(z)=0$ that have any (not necessarily all) complex roots of absolute value 1;

(2) all equations derived from (1) by subjecting $f(z)$ to a non-singular linear transformation $z=(az_1+b)/(cz_1+d)$ whose coefficients are real and rational.

In particular, all irreducible quadratic equations with rational coefficients and negative discriminant belong to this class. The case (1) contains as a subclass all cyclotomic equations of prime degree.

It turns out that if an irreducible equation with rational coefficients has any root on a circle S, or C, the roots are sym-

* A systematic examination of applications of Kronecker's theorem may seem desirable. Compare also D. H. Lehmer's article referred to at the end of the next section.

† Kempner, *On irreducible equations admitting roots of the form* $\alpha+\rho e^{i\theta}, \cdots$, Tôhoku Mathematical Journal, vol. 13 (1918), pp. 253–265.

metrically distributed with respect to this circle. In this statement, two points are called symmetric with respect to a circle when each is the image of the other under transformation by reciprocal radii. This represents a natural extension of the symmetry property of the axis of reals for all equations with real coefficients.

THEOREM 3. *When an irreducible equation has a complex root of the form* $\alpha+\rho e^{i\theta}$, *where* α *and* ρ^2 *are rational and where* ρ *is positive, but may be either rational or irrational, the roots of the equation are distributed in the following manner*:

(1) *Besides* $\alpha+\rho e^{\pm i\theta}$ *there may be other pairs of complex roots with the same* α *and the same* ρ, *that is, of the form* $\alpha+\rho e^{\pm i\theta_1}$.

(2) *If* σ *is a real root,* ρ^2/σ *is a root.*

(3) *Any complex root not contained in* (1) *may, of course, be written in the form* $\alpha+re^{i\phi}$, *with the same* α. *Then the four numbers* $\alpha+re^{\pm i\phi}$, $\alpha+(\rho^2/r)e^{\pm i\phi}$ *are all contained among the roots of the equation.*

The complete extension of the symmetry property of the axis of reals, insofar as I can see, may be stated as follows. Let $f(z)=z^n+a_1z^{n-1}+\cdots+a_n=0$ be an equation with real, but otherwise arbitrary, coefficients, irreducible in the domain $R(a_1, a_2, \cdots, a_n)$, and let us assume that $f(z)=0$ has a root of the form $\alpha+\rho e^{i\theta}$, where α and ρ^2 are both rationally expressible in terms of $a_1, \cdots, a_n$ with rational numerical coefficients. Then all roots of $f(z)=0$ are distributed in the plane of complex numbers symmetrically to the axis of reals and to the circle $\alpha+\rho e^{i\theta}$. We may state also the following theorems.

THEOREM 4. *When an irreducible equation* (*in* $R(1)$) *with real coefficients has any complex root on any circle C, or on any straight line parallel to the axis of imaginaries and at a rational distance from it, the roots are symmetrically distributed with respect to this circle or straight line.*

THEOREM 5. *The points of intersection of two circles S, which form a set that is everywhere dense in the complex plane, are exactly the set of roots of the class of irreducible equations* $az^2+bz+c=0$, *where a, b, and c are rational and* $b^2-4ac<0$. *In other words, the set of points of intersection of S are exactly the set of all complex*

irrationalities. The real quadratic irrationalities are given by the points of the axis of reals belonging to C, but not to S.

It is easily shown that the complex (not real) points of the set of intersections of circles of the set C are the same as the complex (not real) points of the set of intersections of the set S. Through each point of intersection of two circles S (or C) pass an infinite number of circles S (or C).

Therefore, any imaginary quadratic irrationality may be represented in an infinite number of ways by the form $\alpha+\rho e^{i\theta}$, where α and ρ are rational, but any root of an irreducible equation of degree $n>2$ can be represented in at most one way in this form. Consequently, for $n>2$, the equation $\alpha_1+\rho_1 e^{i\theta_1}=\alpha_2+\rho_2 e^{i\theta_2}$, where α_1, α_2, ρ_1, and ρ_2 are rational and where ρ_1 and ρ_2 are positive, implies that $\alpha_1=\alpha_2$ and that $\rho_1=\rho_2$.

It follows from the preceding theorems that if an irreducible equation of degree $n>2$ has roots on a circle S, these roots cannot be among the points of intersection of circles S.

THEOREM 6. *An irreducible equation cannot have roots on more than one circle S unless the circles intersect. On every circle S lie roots of irreducible equations of degree $n>2$, some of whose other roots lie on intersecting circles.*

THEOREM 7. *An irreducible equation $f(z)=0$ can have two roots $\alpha_1+\rho_1 e^{i\theta_1}$, $\alpha_2+\rho_2 e^{i\theta_2}$, where α_1, α_2, ρ_1, and ρ_2 are all rational and $\alpha_1\neq\alpha_2$, only when $|\rho_1-\rho_2|<|\alpha_1-\alpha_2|<\rho_1+\rho_2$.*

THEOREM 8. *When an irreducible equation $f(z)=0$ has complex roots of the form $z=\alpha+\rho e^{i\theta}$, where α and ρ are rational, α and ρ may be found by a finite number of rational operations and the extraction of a square root. Therefore, if such an equation has a root on any circle S, this circle, but not necessarily the root itself, can be constructed by ruler and compasses.*

Let us give a numerical illustration. The irreducible equation $16z^4+72z^3+96z^2+60z+15=0$ is found to have roots of the form $\alpha+\rho e^{i\theta}$, where $\alpha=-1/2, \rho=1/2$. The equation has two real roots, $(-(21)^{1/2}-9\pm[30+10(21)^{1/2}]^{1/2})/8$, and on the circle $-1/2+e^{i\theta}/2$, two complex roots, $\{(21)^{1/2}-9\pm[30-10(21)^{1/2}]^{1/2}\}/8$.

Necessary and sufficient conditions for an irreducible equation to have complex roots of the form $\alpha+\rho e^{i\theta}$, where α and ρ

are both rational, are obtained by extending our theorem on necessary and sufficient conditions that an equation have roots of the form $e^{i\theta}$.

THEOREM 9. *An irreducible equation $f_n(z)=0$ has among its complex roots some roots of the form $\alpha+\rho e^{i\theta}$, where α and ρ are rational, when and only when both of the following conditions are satisfied.*

1. *The system of equations*

$$\rho^n a_0 = f(\alpha), \qquad \rho^{n-2}(na_0\alpha + a_1) = f'(\alpha)$$

has a rational solution (α, ρ).

2. *Letting $z=\alpha+\rho z$ and forming the expressions*

$$g(z_1) = f(\alpha + \rho z_1),$$

$$\phi_{2n}(z_1) = (z_1^2 + 1)^n g\left(\frac{z_1 + i}{z_1 - i}\right) g\left(\frac{z_1 - i}{z_1 + i}\right) = \psi_n(z_1^2) = \psi_n(u),$$

we see that the equation $\psi_n(u)=0$ must admit at least one positive root.

As applications, we shall mention the following cases.

(1) Klein's famous *Modul-figur*, consists, in the plane of complex numbers, of the set of circles of radius 1/2 with centers on the axis of reals at $0, \pm 1, \pm 2, \cdots$, and the system of straight lines parallel to the axes of imaginaries which are tangent to these circles and the system obtained by continued reflection of this figure with respect to itself by reciprocal radii. We may state the following property. *An irreducible equation cannot have roots on more than one circle of the Klein Modul-figur.*

(2) *Every irreducible equation of degree <10 with a root on any circle S is completely solvable by radicals.**

(3) *When an irreducible equation $a_0z^n+a_1z^{n-1}+\cdots+a_n=0$ has a complex root of the form $\alpha+\rho e^{i\theta}$, where α and ρ are rational, the equation $a_0y^n=a_0x^n+a_1x^{n-1}+\cdots+a_n$ has a rational solution (x, y).*

When α and ρ are both integers, the diophantine equation $a_0y^n=a_0x^n+\cdots+a_n$ has an integral solution (x, y).

If the equation $a_0z^n+a_1z^{n-1}+\cdots+a_n=0$ has a complex root

* This can probably be extended to give a result corresponding to the last theorem of §2.

on a circle of rational radius touching the axis of imaginaries at the origin, the equation $a_1x^{n-1}+a_2x^{n-2}+\cdots+a_n=0$ *has a rational root.* The equation $16z^4+72z^3+96z^2+60z+15=0$, which was mentioned above, comes under this heading. The equation $72x^3+96x^2+60x+15=0$ has a root $-1/2$.

In connection with this work I may mention a paper by D. H. Lehmer.* He defines a polynomial as quasi-cyclotomic if all zeros lie on a circle of radius r, with center at the origin, and if the arguments of the roots are commensurable with 2π. His work is connected with our own by his Theorem 3, which is as follows.

Let the roots of $f=0$ *lie on the circle of radius* r. *Then* r *is an integer or the square root of an integer, according as* n *is odd or even.*

An interesting result is his Theorem 6, which reads: *Let the* $n=2k$ *roots of* $f(x)=0$ *be of the form* $(R)^{1/2}\cdot\epsilon_r$, *where* ϵ_r *are roots of unity, and where* R *has no square factor. Then*

$$(R)^{1/2} \leqq \frac{(2k)!}{(k+1)!(k-1)!}.$$

4. *On the Distribution of Complex Algebraic Integers of a Given Degree.* The identification of the points of intersection of the set of circles S (or C) with the totality of complex quadratic irrationalities (integral or otherwise) suggests the problem of investigating the distribution in the plane of complex numbers of *the totality of algebraic integers* $x+iy$ *of a given degree* $n=2$, *or* $n=3$, *and so on.*

The problem appears artificial on account of the fact that the totalities considered do not form domains of rationality. Yet, it seems to possess a degree of interest. For $n=2$, the solution is given practically by inspection. In the $(x+iy)$ plane, these algebraic integers are obviously distributed along the lines $x=k/2$, $(k=0, \pm1, \pm2, \cdots)$. The distribution is the same for all lines $x=0, \pm1, \pm2, \cdots$, and is also the same for all lines $x=\pm1/2, \pm3/2, \cdots$, but it differs in the two sets. The density increases with increasing $|y|$, and so on. In addition, the integral real quadratic irrationalities are of course distributed everywhere densely on the axis of reals.

For the set of all cubic algebraic integers (and eo ipso those

* *Quasi-cyclotomic polynomials*, American Mathematical Monthly, vol. 39 (1932), pp. 383–389.

of higher order), the first suspicion is that they form in the plane of complex numbers an everywhere dense set. Even if this should be so, there is some indication that an examination of the fine structure of the distribution with respect to "gleichmässige Dichtigkeit" may be of interest.

5. *On the Use of the Arguments of the Coefficients for Separating Complex Roots.* In a previous paper,* I emphasized the possibility of using systematically the arguments of the coefficients of an algebraic equation for the separation of the complex roots. Descartes' rule for an upper bound of positive (negative) roots is the first, very special, case of theorems of this kind. I used a very obvious, yet rather powerful, tool. Consider the equation

$$f(z) = a_n z^n + \cdots + a_0 = 0,$$

$$a_\mu = r_\mu e^{i\theta_\mu},\ 0 \leqq \theta_\mu < 2\pi,\ r_\mu > 0;\ z = \rho e^{i\phi},\ 0 \leqq \phi < 2\pi,\ \rho > 0.$$

In the plane of complex numbers, mark from the origin the half-rays or vectors θ_μ. The vector θ_0 is to remain in its original position. The vector θ_1 is to rotate in a positive sense with a constant angular velocity ω, while for $\mu = 1, 2, \cdots, n$, the vector θ_μ rotates with a uniform angular velocity μ times that of θ_1. Vectors θ_μ for which $r_\mu = 0$, that is, for which the coefficient $a_\mu = 0$, are to be ignored. At any moment the vectors give the directions of the vectors representing the terms $a_\mu z^\mu = r_\mu \rho^\mu e^{i(\theta_\mu + \mu\phi)}$. These directions depend only on ϕ and the θ_μ of the coefficients. Most of the theorems are immediate consequences of the fact that the sum of vectors from a common point can certainly not vanish when it is possible to draw a line through the point such that all vectors lie on one side of the line. This fact is used, for example, in the simplest proof of the Gauss-Lucas theorem that all roots of $f'(z) = 0$ lie inside or on the boundary of the smallest convex polygon around the points representing the roots of $f(z) = 0$.

It is, therefore, impossible to have roots of $f(z) = 0$ in any sector (vertex at the origin) $\alpha < \phi < \beta$ for which all vectors lie on one side of a straight line through the origin. Such ϕ-intervals—if they exist—are determined by simple inequalities,

* Kempner, *Über die Separation komplexer Wurzeln*, Hilbert Festschrift, 1922, pp. 49–59 (=Mathematische Annalen, vol. 85 (1922), pp. 49–59).

which in many cases enable us to determine sectors, depending only on the *arguments* of the coefficients, *not* on the *absolute* values, and which are free of roots.

The efficiency of the method is a function of the number of coefficients different from 0, rather than of the degree of the equation. For trinomial equations, a complete separation of the roots is given; for quadrinomial equations, and m-nomial equations, $m>4$, valuable information usually is obtained, even though complete separation may not be accomplished.

The results are indicated by the following illustrations.

(1) *Binomial equations*: $a_n z^n - a_0 = 0$. We obtain n null sectors (half-rays), each of which contains exactly one root, in agreement with elementary algebra.

(2) *Cubic equations*: $z^3 + az + b = 0$, where a and b are real. The three roots have arguments ϕ_1, ϕ_2, ϕ_3 ($\phi = 0$, real positive roots; $\phi = \pi$, real negative roots);

$$a > 0,\ b > 0: \quad \phi_1 = \pi, \quad \frac{\pi}{3} < \phi_2 < \frac{\pi}{2}, \quad \frac{3\pi}{2} < \phi_3 < \frac{5\pi}{3};$$

$$a > 0,\ b < 0: \quad \phi_1 = 0, \quad \frac{\pi}{2} < \phi_2 < \frac{2\pi}{3}, \quad \frac{4\pi}{3} < \phi_3 < \frac{3\pi}{2};$$

$$a < 0,\ b > 0: \quad \phi_1 = \pi, \quad 0 \leqq \phi_2 < \frac{\pi}{3}, \quad \frac{5\pi}{3} < \phi_3 \leqq 2\pi;$$

$$a < 0,\ b < 0: \quad \phi_1 = 0, \quad \frac{2\pi}{3} < \phi_2 \leqq \pi, \quad \pi \leqq \phi_3 < \frac{4\pi}{3}.$$

(3) *Trinomial equations*: $a_n z^n + a_m z^m + a_0 = 0$, $a_n = r_n e^{i\theta_n}$, $a_m = r_m e^{i\theta_m}$, $a_0 = r_0 e^{i\theta_0}$. For such equations, we obtain complete separation of roots, one in each of n sectors.* Special precautions are required for dealing with the cases in which two sectors adjoin and the roots lie on the common boundary line. As an illustration, consider equations of the form

$$a_5 z^5 + a_2 z^2 + a_0 = 0, \quad \theta_5 = 0, \quad \theta_2 = 2\pi/5, \quad \theta_0 = 5\pi/6.$$

Each of the following five sectors contains exactly one root: $66° < \phi_1 < 84°$, $129° < \phi_2 < 138°$, $204° < \phi_3 < 210°$, $282° < \phi_4 < 309°$,

* Partly overlapping results by Nekrassoff, Mathematische Annalen, vol. 29 (1887), pp. 413–440, who has different aims and uses entirely different methods.

$324° < \phi_5 < 354°$. The openings of the sectors are, respectively, 18°, 9°, 6°, 27°, 30°. This result holds, of course, independently of the absolute values of the coefficients.

(4) *Quadrinomial equations.* As an illustration, consider any equation of the form

$$a_{10}z^{10} + a_9z^9 + a_5z^5 + a_0 = 0, \qquad (a_0, a_5, a_9, a_{10} > 0).$$

Roots are separated into ten sectors: $18° < \phi_1 < 36°$, $36° < \phi_2 < 60°$, $90° < \phi_3 < 108°$, $108° < \phi_4 < 140°$, $162° < \phi_5 \leqq 180$, $180° \leqq \phi_6 < 198°$, $\cdots$, $324° < \phi_{10} < 342°$. The largest opening is 32°, the smallest 18°. Complete separation is partly due to the high degree of the equation, despite the small number of terms. For quadrinomial equations, complete separation is no longer guaranteed.

(5) *Mechanical Determination of Sectors.* An instrument of the type of a "planetarium" which permits n vectors to rotate from arbitrary initial positions with angular velocities which are to each other in the ratios $n:n-1:\cdots:3:2:1$ would permit the mechanical determination of sectors free from roots. An ordinary watch will separate all roots of any equation of the form

$$z^{12} \pm az^{\lambda} \pm b = 0, \qquad (a > 0, b > 0, \lambda = 1 \text{ or } 11).$$

Consider, for example, any equation of the form $z^{12}+a_1z+a_0=0$, $(a_0, a_1>0)$. We obtain

$$15° < \phi_1 < 16\frac{4}{11}°, \quad 45° < \phi_2 < 49\frac{1}{11}°, \quad 75° < \phi_3 < 81\frac{9}{11}°,$$

$$105° < \phi_4 < 114\frac{6}{11}°, \quad 135° < \phi_4 < 147\frac{3}{11}°, \quad 165° < \phi_6 \leqq 180°,$$

with a similar distribution of 6 sectors from 180°–360°. The openings of the sectors are, respectively,

$$\frac{4}{11}°, 4\frac{1}{11}°, 6\frac{9}{11}°, 9\frac{6}{11}°, 12\frac{3}{11}°, 15°, 15°, \cdots, 1\frac{4}{11}°.$$

N. Obreschkoff* has given another extension of Descartes' rule. His result is as follows.

---

* *Sur un problème de Laguerre,* Comptes Rendus, vol. 177 (1923), pp. 102–104.

Let $f(z)=a_0+a_1z+\cdots+a_nz^n=0$ *be an equation with real coefficients. The number* v_a *of variations in the sequence of its coefficients is not only an upper bound for the number* N *of its positive roots, but also an upper bound for the number of roots lying in the sector* $|\arg z|<\pi/n$. *If the number of roots lying in the sector is less than* v_a, *it is smaller by an even number.*

I. J. Schoenberg was kind enough to let me see the manuscript of an article soon to be published in the Annals of Mathematics. In the first part of this paper, the author introduces an ingenious extension of the notion of a sequence of variations in signs to the case of complex coefficients. This is accomplished by the so-called *separating double-sectors* for the coefficient a_ν. To define these, we imagine in the plane of complex numbers the coefficients a_ν marked by their corresponding points. Through the origin draw any two straight lines l_1 and l_2 that divide the plane into four consecutive sectors A_1, A_2, A_3, A_4, with the following property: one of two pairs of opposite sectors, for example, A_1 and A_3, shall not contain any of the points a_ν in their interior. If the common opening of A_1 and A_3 is ψ, $(0<\psi\leqq\pi)$, we call A_1 and A_3 together a *separating double-sector* (of aperture ψ) for the coefficients a_ν. In going in order through the sequence of points $a_0, a_1, \cdots, a_n$, we count the number of times we have to pass from the sector A_2 to A_4 or vice-versa. This number is called $v_a(S)$ and is equal to the number of variations of the sequence $a_0, \cdots, a_n$ with respect to the separating double-sector $S=(A_1, A_3)$. It is evident that for all a_ν real, $v_a(S)$ (taking $\psi=\pi$) agrees with the ordinary *number of variations*. Schoenberg then derives the following theorems.

Let $f(z)=a_0+a_1z+\cdots+a_nz^n=0$, *where* a_ν *is real or complex. In the plane of complex numbers, mark the points* a_ν *and draw a separating double-sector of aperture* ψ, $(0<\psi\leqq\pi)$, *as can always be done. Then the number of roots lying in the sector* $|\arg z|<\psi/n$ *is either equal to* $v_a(S)$, *or is less by an even number. To find an upper bound for the number of roots of* $f(z)=0$ *lying in a sector whose bisectrix makes with the positive axis the angle* θ, *a substitution* $z=e^{i\theta}z'$ *is carried out.*

For any sector with the vertex at the origin and of aperture $2\pi/[n(n+1)]$ *an upper bound of the number of roots lying in the sector can be established.*

The rest of this interesting paper deals with applications of

these ideas to equations of various types which fall outside the frame of this paper.

6. *Absolute Value, Argument, Real Part, Pure Imaginary Part of Complex Roots.*

Given

$$f(z) = a_0z^n + a_1z^{n-1} + \cdots + a_n = 0,$$

where $z=\alpha+i\beta=re^{i\phi}$ is a root. It is well known that, if a_ν are integers, then r, $\cos\phi$, $\sin\phi$, α, and β all are algebraic numbers, and satisfy equations with integral coefficients which may be derived from $f(z)=0$ by rational operations. It would be of some interest to have a systematic examination of these equations and their use for the determination of complex roots. At present the information is badly scattered in the literature.*

Difficulties arise in all cases: (1) because the equation for any of r, $\cos\phi$, $\sin\phi$, α, β is of high degree, usually $n(n-1)$, with reductions in some cases for real rational a_ν: for example, for $a_0z^2+a_1z+a_2=0$, $z=\alpha+i\beta$, $\beta\neq 0$, β satisfies a second degree equation, but α satisfies a linear equation, with integral coefficients; (2) because the equations necessarily have a large number of extraneous roots, which are sometimes not readily distinguishable from the roots required. Some aspects of a systematic examination and classification of methods of solution along these lines are contained in two master's theses.†

It is interesting to note that the equations for r and for $e^{i\phi}$ are related in structure (for a_ν real). Let $z=re^{i\theta}=rs$ be a complex root of $f(z)=0$.

Then r/s is also a root, and hence

$$\begin{aligned} a_0r^ns^n + a_1r^{n-1}s^{n-1} + \cdots + a_{n-1}rs \quad + a_n \quad &= 0, \\ a_0r^n + a_1r^{n-1}s \quad + \cdots + a_{n-1}rs^{n-1} + a_ns^n &= 0. \end{aligned}$$

Eliminating, we have as equations for r and s, respectively,

* References are too numerous to mention. We refer to just one interesting paper by Hayashi, Tôhoku Mathematical Journal, vol. 3 (1913), pp. 110–115, for the derivation of the equation for the absolute values. The degree of the resulting equation is incorrectly given [n, instead of $n(n-1)$].

† R. C. Huffer, University of Illinois, 1920, and A. W. Randall, University of Colorado, 1929.

$$\begin{vmatrix} a_0r^n & a_1r^{n-1} \cdots a_n & & \\ & a_0r^n \quad \cdots \quad a_n & & \\ \cdots & \cdots & & \\ a_n & a_{n-1}r \cdots a_0r^n & & \\ & a_n \quad \cdots \quad a_0r^n & & \\ \cdots & \cdots & & \end{vmatrix} = 0, \qquad \begin{vmatrix} a_0s^n & a_1s^{n-1} \cdots a_n & & \\ & a_0s^n \quad \cdots \quad a_n & & \\ \cdots & \cdots & & \\ a_0 & a_1s \quad \cdots a_ns^n & & \\ & a_0 \quad \cdots \quad a_ns^n & & \\ \cdots & \cdots & & \end{vmatrix} = 0.$$

In connection with the problem of numerical solution of algebraic equations, I should like to mention a doctoral thesis* by A. J. Lewis on the solution of algebraic equations by infinite series, since it has not yet been published. As in practically all theoretical work on solutions by series, the background is given by Lagrange's expansion. The method was elaborated by Nekrassoff (1887), McClintock (1894), Lambert (1908), and others. However, the whole work is very fragmentary in character; and Lambert's work, in particular, is open to criticism. Lewis has given a more detailed treatment of the expansions than was contained in the literature, has corrected Lambert's errors, and has made a step in advance in the investigation of the convergence conditions. Trinomial and quadrinomial equations are discussed in detail, and the treatment for equations with more than four terms is indicated. The formulas are very complex, but this is due to the intrinsic difficulties of the problem.

7. *On a Graphical Method for the Separation and Computation of Complex Roots.* The next article I shall take the liberty of reporting more in detail, since it is published in a journal not easily available to mathematicians.† In this paper, three methods of graphical representation are compared.‡

(1) The representation familiar from conformal mapping, or from proofs of the Fundamental Theorem of Algebra: $f(z) = f(x+iy) = u(x, y) + iv(x, y)$. The intersections of the two curves $u(x, y) = 0$, $v(x, y) = 0$ in the xy plane give the roots $x+iy$ of $f(z) = 0$. In order to find the number of roots of absolute value ρ, where ρ is arbitrarily assigned, it is only necessary to

* A. J. Lewis, Thesis, University of Colorado, 1932.

† Kempner, *On the separation and computation of complex roots of algebraic equations*, The University of Colorado Studies, vol. 16 (1928), pp. 75–87.

‡ For the careful computation of all graphs of this paper, I am under great obligation to my colleague Professor C. A. Hutchinson.

find the number of points of intersection of the u curve and the v curve inside a circle about the origin of radius ρ.

It is possible, by means of Kronecker's *characteristics*, or by equivalent methods, to determine from the figure the number of roots of $f(z)=0$ inside any closed contour by considering nothing but the points on the contour where the curves $u=0$ and $v=0$ enter and leave the region. We shall use this idea later.

(2) Let

$$f(z) = c_n z^n + c_{n-1}z^{n-1} + \cdots + c_0 = 0,$$

where $c_k=a_k+ib_k=r_k(\cos\gamma_k+i\sin\gamma_k)=r_ke^{i\gamma_k}$, and $z=x+iy=\rho(\cos\phi+i\sin\phi)=\rho e^{i\phi}$.

We plot

$$u = \sum_{k=1}^{n} r_k\rho^k\cos(k\phi+\gamma_k), \qquad v = \sum_{k=1}^{n} r_k\rho^k\sin(k\phi+\gamma_k)$$

in the following manner. For ρ, the absolute value of z, we select an assigned value which is held fixed. Then u and v become functions of ϕ alone, say $u=\psi_1(\phi)$, $v=\psi_2(\phi)$. Thus u and v may be considered as coordinates of a curve given in parametric form, with ϕ as the parameter. We plot the curve in a uv system of coordinates, choosing the horizontal axis as the u axis and the vertical axis as the v axis.

(3) We now plot the two curves $u=\psi_1(\phi)$, $v=\psi_2(\phi)$ separately, but both in the same system of coordinates. For both curves, we take the horizontal axis as the ϕ axis, the vertical axis as the u axis for the first curve, and as v axis for the second curve.

Discussion of (2). To a circle in the complex z plane about the origin, of given radius ρ, corresponds in the uv plane, which we may also interpret as the $(u+iv)$ plane, a certain closed curve. If N is the number of roots of $\psi(z)=0$ in the circle, and if we assume that no roots lie on the circumference, each root being counted with its proper multiplicity, then $2\pi N$ represents, as is known from the theory of analytic functions of a complex variable, the change of amplitude when the curve is once completely traced in the $(u+iv)$ plane.

Simple considerations of a type familiar from the theory of characteristics will permit us to read off this number N from our representation (2). To find the number of roots in a circle of radius ρ about the origin in the z plane, we consider the corresponding closed curve (2). Trace the curve (in either the posi-

tive or negative sense), starting from an arbitrarily chosen point, and mark as points β all points where the curve crosses the u axis ($v=0$), and as points α all points where it crosses the v axis ($u=0$). Now we determine the change in the argument between two consecutive points of intersection with the u axis, as we move along the curve. We call this argument ψ, since it is not necessarily the same as the parameter ϕ, the argument of the construction (3). The simplest case occurs when our points α and β form an alternating sequence, $\alpha\beta\alpha\beta\cdots\alpha\beta$. To a change from any α to the next α corresponds each time a change of π in ψ. All these changes are in the same sense; that is, ψ is either each time increased by π or each time decreased by π. Hence ψ increases or decreases by $m\pi$, where m is the number of β, and there are therefore $m/2$ roots in the interior of the circle of radius ρ about the origin in the $(x+iy)$ plane. But the number $m=2N$ is correctly expressed by the following theorem for the special case $\epsilon_1=\epsilon_2=\cdots=\epsilon_\nu=1$, $\nu=m$.

THEOREM 10. *Denote by $\beta_1, \beta_2, \cdots$ the points of intersection of curve (2) with the u axis, by $\alpha_1, \alpha_2, \cdots$ the points of intersection with the v axis, and consider the chain of the points α and β cyclically closed. If then, starting from any α, the number of β points between two consecutive α points is, respectively, $\epsilon_1, \epsilon_2, \cdots, \epsilon_k$, then the number of roots z of $f(z)=0$ for which $|z|<\rho$ is given by the formula*

$$2N = \left| \sum_{k=1}^{\nu} (-1)^{(\epsilon_1+1)+(\epsilon_2+1)+\cdots+(\epsilon_k+1)} \right|.$$

I omit the proof that this formula, derived for the special case of alternating α and β, is true for any sequence of the α and β. These considerations lead to the following rule.

RULE. *Write down, in proper order, the numbers which represent the number of β in the successive α-intervals. Under the first of these numbers write $+1$, if it is odd, or -1, if it is even. Write then under each odd number that one of the two numbers, $+1$ or -1, which is already written under the last preceding number; but under each even number write $+1$ when -1 is written under the last preceding number, and write -1 when $+1$ was written under the preceding number. The sum of all of these numbers $+1$, -1 is necessarily even, and the absolute value of the sum is $2N$.*

Discussion of (3). From the constructions (2) and (3), we find the following relation between them.

It is clear that to every curve (3) will correspond uniquely a curve (2), since for every value of ϕ both u and v are uniquely determined. On the other hand, from a curve (2) a curve (3) is not uniquely determined: according to the manner of parametric representation of u and v by means of ϕ, an infinite number of curves (3) may be obtained. All curves (3), however, which correspond to the same curve (2), are related to one another in a simple fashion.

The theorem is applied, without modification, to (3). In our curves (1), drawn for a given value ρ, and with a common ϕ axis, the intersections of the u curve with the ϕ axis and the intersections of the v curve with the ϕ axis correspond in a definite manner to the intersections of the uv curve in (2) with the u axis and the v axis. To make this clear, it is only necessary to consider carefully a special case. Consider the uv curve treated in Fig. 1_2.* To this figure corresponds (schematically) Fig. 2_2 as is immediately checked. In comparing these two figures we have only to keep in mind that every time the curve (2) crosses the u axis, we have $v=0$, so that in (3) the v curve crosses the ϕ axis, and that every time the curve (2) crosses the v axis, we have $u=0$, so that in (3), the u curve crosses the ϕ axis, and finally, that the combined sequence of the points of intersection of the u curve and the v curve with the ϕ axis follow in the same order as the $\alpha\beta$ sequence in (2).

It follows that we get exactly the same $\alpha\beta$ sequence, whether we read it off from (2) or from the sequence obtained by marking in (3) as α all points of intersection of the u curve with the ϕ axis, and as β all points of intersection of the v curve with the ϕ axis. We have, therefore, the following theorem.

THEOREM 11. *Plot, as explained under* (3), *the curves*

$$u = \sum_{k=0}^{n} r_k \rho^k \cos(k\phi + \gamma_k), \qquad v = \sum_{k=0}^{n} r_k \rho^k \sin(k\phi + \gamma_k),$$

for a given ρ. *Name the points of intersection of the u-curve with the* ϕ *axis,* $\alpha_1, \alpha_2, \cdots$, *and name the points of intersection of the v curve with the* ϕ *axis,* $\beta_1, \beta_2, \cdots$. *Consider the combined se-*

* The index 2 in 1_2 indicates that the figure refers to construction (2), etc.

quence of the α and β points in their natural order of magnitude. Assume the sequence closed cyclically, and let $\epsilon_1, \epsilon_2, \cdots, \epsilon_r$, respectively, be the number of β points between two consecutive α points; then the number N of roots of $f(z)=0$ for which $|z|<\rho$ is given by the formula

$$2N = \left| \sum_{k=1}^{r} (-1)^{(\epsilon_1+1)+(\epsilon_2+1)+\cdots+(\epsilon_k+1)} \right|.$$

This number may be evaluated by the rule stated under Theorem 10.

EXAMPLE. The following example is only schematic, and shows how the rule and the theorem are applied to more complicated cases.

Figures 1_2 and 2_3 are corresponding figures. From either of them we read off the $\alpha\beta$ sequence, and apply the rule or the theorem.

	β	α	α	β	α	α	α	β	β	α	β	α	β	α	β	β
$\epsilon_k=$	0			1	0	0		2			1		1		3	
	-1			-1	$+1$	-1		$+1$			$+1$		$+1$		$+1$	

Using the rule, we find $2N=2$, $N=1$.

Using the theorem, we find

$$\epsilon_k+1 = \quad 1 \quad 2 \quad 1 \quad 1 \quad 3 \quad 2 \quad 2 \quad 4$$

$$2N=(-1)^1+(-1)^3+(-1)^4+(-1)^5+(-1)^8+(-1)^{10}+(-1)^{12}$$
$$+(-1)^{16}=2, \qquad N=1.$$

If we had started with β intervals and had considered the number of α in each β interval, we should have found

	β	α	α	β	α	α	α	β	β	α	β	α	β	α	β	β
$\epsilon_k=$	2			3			0	1		1		1		0	0	
$\epsilon_k+1=$	3			4			1	2		2		2		1	1	

$$2N=(-1)^3+(-1)^7+(-1)^8+(-1)^{10}+(-1)^{12}+(-1)^{14}+(-1)^{15}$$
$$+(-1)^{16}=2, \quad N=1.$$

Application of Theorem 11. We make use of the theorem to isolate the complex roots of an equation in concentric rings ($\rho_1<\rho<\rho_2$) around the origin as center. The procedure is as follows. For two given positive values of ρ, $\rho_2>\rho_1$, we assume

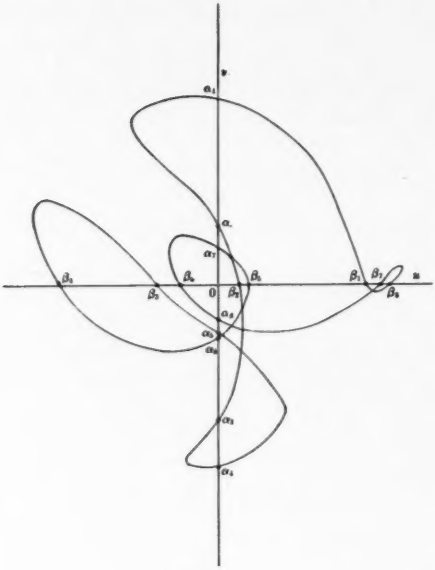

FIG. 1₂

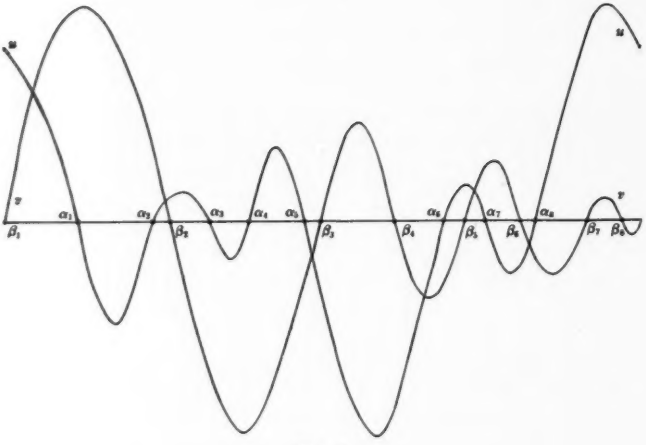

FIG. 2₂

that the curves $u(\phi)$ and $v(\phi)$ are drawn in our system (3). By Theorem 10 we read off the numbers $N(\rho_1)$, $N(\rho_2)$ of roots of absolute value smaller than ρ_1 and ρ_2, respectively. Then $N(\rho_2) - N(\rho_1)$ gives the exact number of roots in the ring-region $\rho_2 > \rho > \rho_1$. Here we assume, since this case is easily disposed of, that no root lies on the circumference of either of the circles ρ_1 or ρ_2.

In case we have the same sequence of α and β for ρ_1 as we have for ρ_2, there are obviously no roots in the ring-shaped region $\rho_1 < |z| < \rho_2$.

Since there exist well known simple upper and lower bounds for the absolute value of the roots of a given equation, we may start out with two such values for ρ, and, by appropriate choice of new values of ρ, we can separate our roots according to their absolute value.

It is a simple matter to plot curves of the kind used here, particularly because we are really interested only in the se-

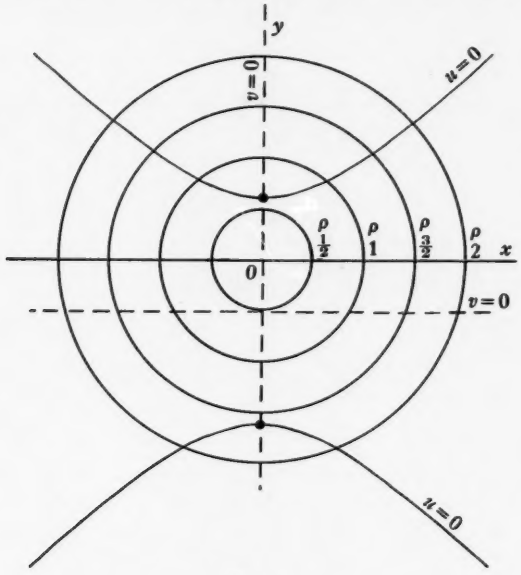

FIG. 3$_1$

quence of the points of intersection of the u curve and the v curve with the ϕ axis. The chief advantage of the present method may lie in the fact that instruments exist with which curves of type (3) can be drawn. Such an instrument is the Michelson-Stratton harmonic analyzer, for which an accuracy of about 1 per cent to 2 per cent is claimed.

EXAMPLE. Consider the equation

$$z^2 + iz + 1 = 0, u = \rho^2 \cos 2\phi - \rho \sin \phi + 1, v = \rho^2 \sin 2\phi + \rho \cos \phi.$$

Figure 3_1 gives the classical construction, which may be used for the verification of our results. Figure 4_2 gives, for $\rho = 1/2$, the parametric representation of $u = u(\phi)$, $v = v(\phi)$ in a uv-system of coordinates, while Fig. 5_3 gives u and v both as functions of ϕ. Figures 6_2 and 7_3 are the corresponding curves for $\rho = 1$; Figs. 8_2 and 9_3 are the curves for $\rho = 3/2$; and finally Figs. 10_2 and 11_3 are the curves for $\rho = 2$.

For $\rho = 1/2$, from Fig. 5_3, we have

$$\begin{matrix} & \beta & \beta \\ \epsilon_k = & 0 & 0, \end{matrix} \qquad 2N = (-1)^1 + (-2)^2 = 0, \qquad N = 0.$$

This is obvious, since the u curve does not intersect the ϕ axis at all. For $\rho = 1$, from 7_3, we find

$$\begin{matrix} & \alpha & \beta & \alpha & \beta & \beta & \beta \\ \epsilon_k = & 1 & & 3 & & , \end{matrix} \quad 2N = (-1)^2 + (-1)^4 = 2, \quad N = 1.$$

For $\rho = 3/2$, from 9_3, we get

$$\alpha \ \beta \ \alpha \ \beta \ \beta \ \beta, \quad N = 1, \text{ as for } \rho = 1.$$

For $\rho = 2$, from 11_3, we obtain

$$\begin{matrix} & \alpha & \beta & \alpha & \beta & \alpha & \beta & \alpha & \beta \\ \epsilon_k = & 1 & & 1 & & 1 & & 1 & \end{matrix}$$

$$2N = (-1)^2 + (-1)^4 + (-1)^6 + (-1)^8 = 4, \qquad N = 2.$$

Therefore, the equation $z^2 + iz + 1 = 0$ has one root in the ring-region $1/2 < \rho < 1$ and one in the region $3/2 < \rho < 2$. This agrees with Fig. 3_1. The absolute values of the two roots are $5^{1/2}/2 - 1/2 = 0.6+$ and $5^{1/2}/2 + 1/2 = 1.6+$, respectively.

It will be recognized immediately that the graphs shown in

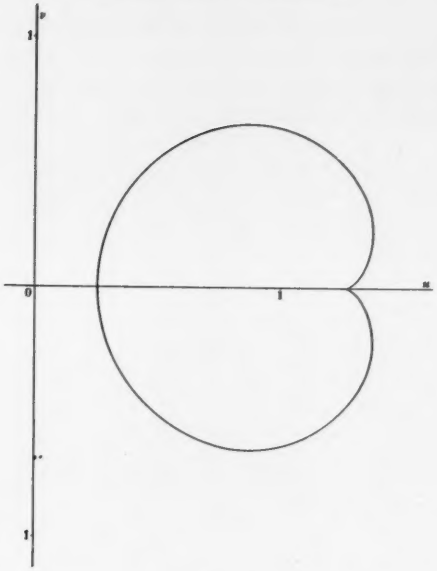

FIG. 4₂

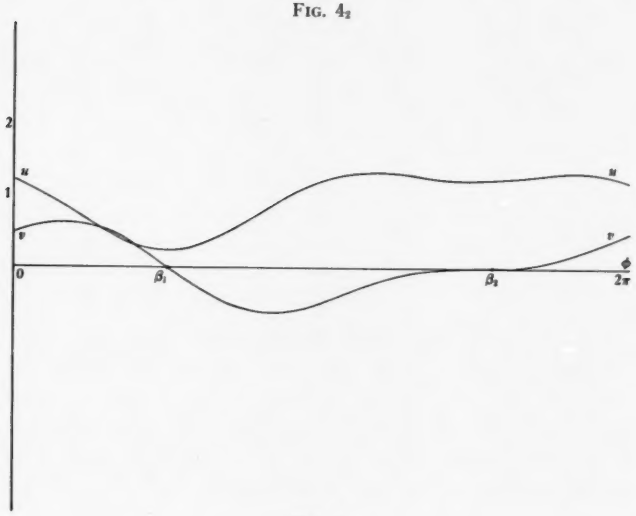

FIG. 5₂

Figs. 5_2, 7_3, 9_3, 11_2 also give information concerning the *arguments* of the complex roots.

8. *Separation of Roots in Sectors.* Without insisting at the present time, I shall only indicate how these same ideas can be utilized to separate roots in sectors from the origin by interchanging the role of r and ϕ in the former process.

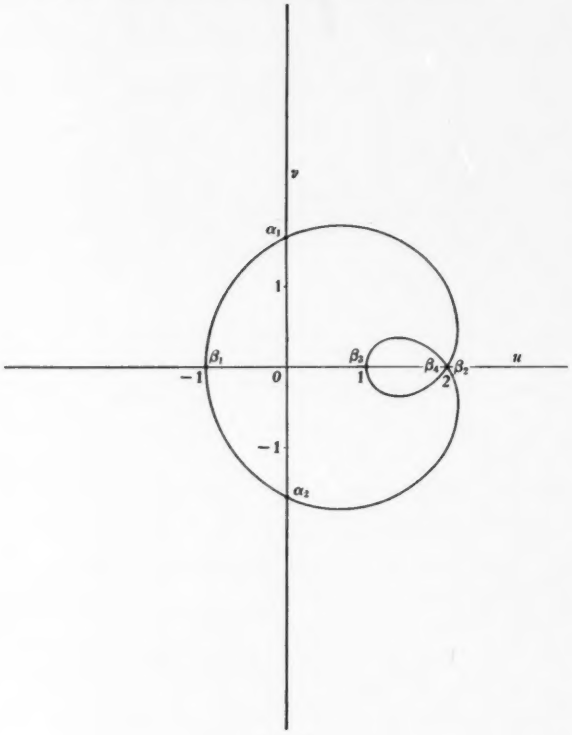

FIG. 6_2

Given the equation $f(z)=a_0z^n+\cdots+a_n=0$, where a_ν is real or complex, and again $z=\rho e^{i\phi}=\rho(\cos\phi+i\sin\phi)$. Now, for a given ϕ, we have $u=u(\rho)$, $v=v(\rho)$. Plot $u(\rho)$ and $v(\rho)$ in the same system of coordinates, with ρ as horizontal axis, and with u

and v as vertical axes. We know how to find upper and lower bounds ρ_1 and ρ_2 for the absolute value of the roots. The value of ρ (always >0) is to be restricted to the interval between these upper and lower bounds. If the sequence of α and β is the same in the interval (ρ_1, ρ_2) for $\phi=\phi_1$ and for $\phi=\phi_2$, there are no roots of $f(z)=0$ in the sector $\phi_1 \leqq \phi \leqq \phi_2$. If the sequence of α and β is not the same for $\phi=\phi_1$ and $\phi=\phi_2$, there are roots in the sector and the exact number may be read off by the same kind

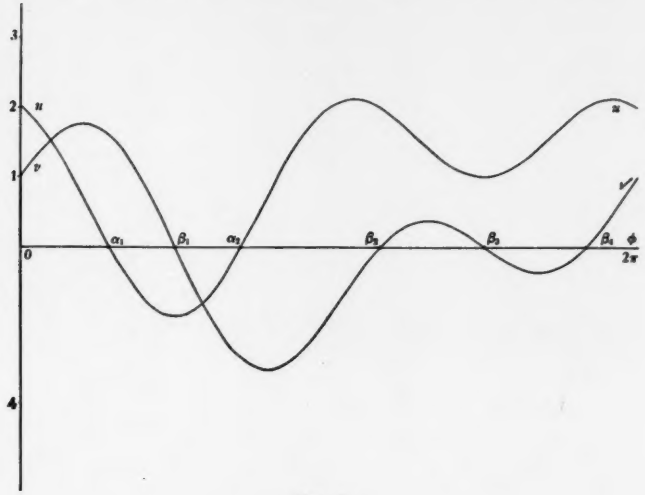

FIG. 7$_3$

of theorem or rule as that which was derived above. The details of this are reserved for another occasion. At present, we restrict ourselves to the verification that $z^3-2z-2=0$ has a root in the sector $3\pi/4 \leqq \phi \leqq \pi$, with $1<\rho<3/2$:

$$z^3 - 2z - 2 = 0.$$

$$u = \rho^3 \cos 3\phi - 2\rho \cos \phi - 2, \quad v = \rho^3 \sin 3\phi - 2\rho \sin \phi.$$

$$\phi = 3\pi/4: \quad u = \frac{1}{2}(2)^{1/2}\rho^3 + (2)^{1/2}\rho - 2, \quad v = \frac{1}{2}(2)^{1/2}\rho^3 - (2)^{1/2}\rho;$$

$$\phi = \pi: \quad u = -\rho^3 + 2\rho - 2, \quad v = \rho^3 \sin \pi - 2\rho \sin \pi.$$

Since for $\phi=\pi$, v reduces to 0 identically, we select $\phi=\pi-\epsilon$, where ϵ denotes some small positive angle. Then, from the relation $\lim_{\epsilon\to 0}[(\sin\epsilon)/\epsilon]=1$, we find $v=3\rho^3\epsilon-2\rho\epsilon$, approximately.

For $\phi=3\pi/4$, our u curve, for $1<\rho<3/2$, lies entirely above the ρ axis, and the v curve crosses the ρ axis once (no α, one β).

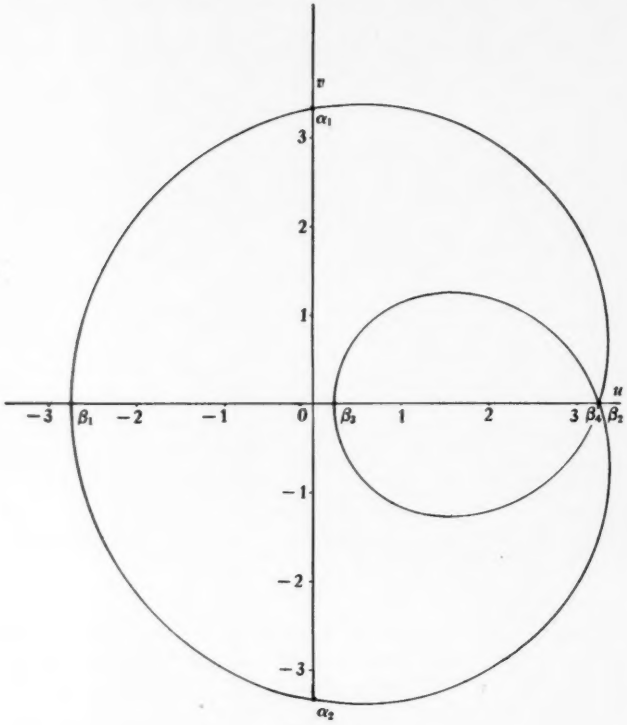

FIG. 8₂

For $\phi=\pi-\epsilon$, neither the u curve nor the v curve cut the ρ axis for $1<\rho<3/2$ (no α, no β). There must be roots of $f(z)=0$ that lie in the sector $3\pi/4\leqq\phi\leqq\pi$.

We do not have, as we did before, an instrument available for drawing the curves, but, on the other hand, they are of par-

ticularly simple type, namely, polynomial curves. Besides we need, of course, not the curves themselves, but only their (approximate) intersections with the ρ axis.

9. *Polynomial Curves and Complex Roots.* I shall discuss one last topic also with some degree of completeness, in order to piece together ideas which are scattered and more frequently implied than explicitly announced in the literature.

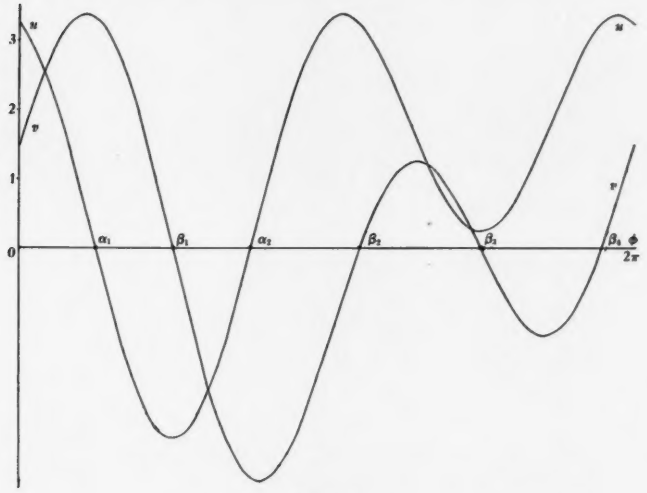

FIG. 9₃

The importance of polynomial curves for the *real* roots of algebraic equations is fundamental and obvious. But the systematic use of graphs for the purpose of locating complex roots of equations of the form $f(z)=a_0z^n+\cdots+a_n=0$, where a_ν is assumed, for simplicity only, to be real, is not so widely known as it should be, except for small values of n, such as occur in the problem of the location of complex roots of quadratic or cubic equations with real coefficients.

The character of the real graph of $f(z)$ is of great influence in some of the graphical methods for complex roots. I have classified real polynomial curves of any given degree n according to

their geometric behavior, and have given existence proofs for the various types.*

Of particular importance for our immediate purpose are certain considerations which permit us to represent in a natural

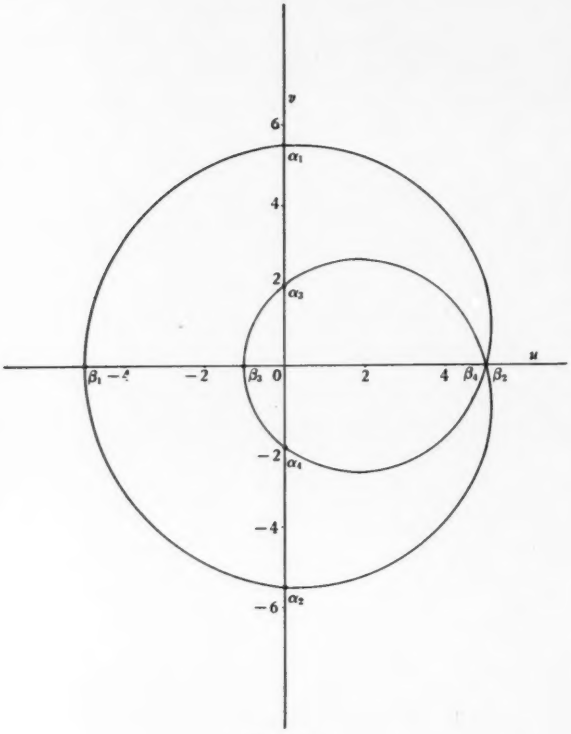

FIG. 10_2

manner some parts of the complex graph, and which, although the germ of the method goes back at least to Poinsot and his

* Kempner, *On the shape of polynomial curves*, I. Tôhoku Mathematical Journal, vol. 37 (1933), pp. 347–362; Part II is to appear in the same journal. Compare also American Mathematical Monthly, vol. 40 (1933), pp. 469–470.

cordes idéaux, are not given the place they deserve in modern presentations.*

The simple idea consists in an attempt to make the best of a bad situation. Since it is impossible to give in a three-dimen-

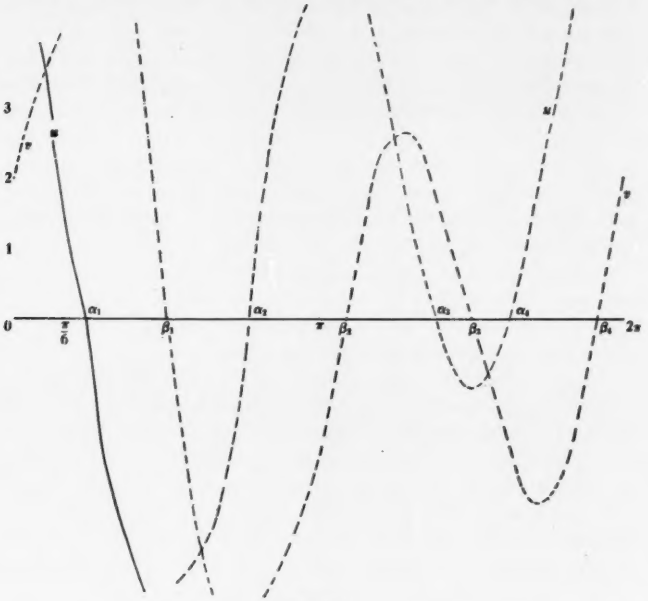

FIG. 11$_3$

sional space a simple satisfactory representation of a function $w=f(z)$ of a complex variable, we try to save as much as we can. We may use our three dimensions as follows. Consider only real values of $f(z)$, but permit z to assume *all* values, real and complex, which make $f(z)$ real.

Letting $f(z)=f(x+iy)=u(x,y)+iv(x,y)=u+iv$, we have,

* Phillips and Beebe, *Graphical Algebra*, 2nd ed. (1887) (out of print), make much use of it in an elementary study of equations of the second, third, and fourth degrees with real coefficients, but with no attempt at a systematic treatment. In particular, the curves $u(x, y)=0$, $v(x, y)=0$ in the xy plane are not considered. The space curves in our text are to be found in Phillips and Beebe.

for $f(z)$ real, $v(x, y)=0$, $f(z)=u(x, y)$. If we use a horizontal xy plane for $z=x+iy$, and a vertical axis for $u(=f(z)$ real), we shall obtain a space curve in our (xyu)-system of coordinates. The projection of the space curve onto the xy plane consists exactly of the curve $v(x, y)=0$, and the functional value is the corresponding value of $u(x, y)$.

It is evident at once that this representation is related to the theory of functions of a complex variable, since the space curve is also given as the curve of contact between the surface of absolute values $|f(z)|$ and the surface of real values $u(x, y)$ for $u(x, y)>0$, and as the curve of contact between the surfaces $-|f(z)|$ and $u(x, y)<0$. Again, the space curves may be interpreted as the intersections of the surface $u(x, y)$ with the cylindrical surface $v(x, y)=0$.

These space curves serve as a skeleton for the function. They are important for us, (a) because we can study their general character with great ease, (b) because by means of them we can always represent all roots of the equation.

(1) For a real function $f(z)=a_0z^n+\cdots+a_n$ with maxima and minima of the first order ($f'=0$, $f''\neq 0$) we obtain a space curve of the following type. First we have in the xu plane the ordinary real graph, but in addition we have attached to each extreme a *handle*, symmetric to the plane of the real graph and penetrating it under a right angle. These handles always extend to infinity and their projections in the xy plane are (as stated) branches of the curve $v(x, y)=0$, the properties of which are well known, sometimes under the name of Lucas' *stelloids*. In certain cases such a handle lies entirely in a plane, which then is for reasons of symmetry parallel to the yu plane. This happens, for example, for the real quadratic function* az^2+bz+c, in agreement with the fact that $v(x, y)=0$ now consists of a line parallel to the y axis, besides the x axis. Similarly, for real polynomial curves with a line of symmetry, which then necessarily have a maximum or a minimum on this line, and which we assume to be an extreme of the first order, the handle attached at that extreme lies entirely in the plane perpendicular

* Leading to one of the classical constructions of complex roots of quadratic equations with real coefficients.

to the xu plane and contains the line of symmetry. For example,* consider the function $f(z)=z^4-2z^2=(z-(2)^{1/2})(z+(2)^{1/2})z^2$. Here $v(x, y)=0$ consists of the curves $y=0$, $x=0$, $x^2-y^2-1=0$. The projection of the real part of the curve is $y=0$; the line $x=0$ contains the handle attached at the maximum; the branches of the hyperbola are the projections of the two handles attached at the minima. These last two handles therefore do not lie in planes. The distance of their points from the xy plane increases monotonically as the projection moves along the hyperbola to infinity.†

The same considerations may be applied to transcendental functions. For $\sin z$, the result is particularly simple: $v(x, y)=0$ consists in the xy plane of the axis of reals (the x axis) and of the equidistant lines $x=\pm(2k+1)\pi/2$, $(k=0, 1, 2, \cdots)$. Hence, all handles lie in planes perpendicular to the xu plane and parallel to the y axis. The handles are nothing but catenaries, attached at the maxima and minima, and extending upward (downward) from a maximum (minimum). Our later remarks concerning the utilization of these extended graphs for the location of complex roots of an equation $a_0z^n+\cdots+a_n=0$ are immediately applicable to the location of complex roots of $\sin z=c$, (c real, $|c|>1$). Applied to any two real circles in a plane, the handles lead to a neat interpretation of the circular points.

For any equation of the form $a_0z^n+\cdots+a_n=0$ with real coefficients and with the maximum number $n-1$ of real extremes (for example, $f(z)=(z-\alpha_1)\cdots(z-\alpha_n)+\beta$, $\alpha_1\neq\alpha_2\neq\cdots\neq\alpha_n$ real, β real), the structure of the real curve and all of its handles is of such simple character that it seems a pity not to make use of it for approximating the location of complex roots. *With but little practice, it is a simple matter to estimate roughly the points where this space curve intersects any plane parallel to the xy plane. These points, read off in the form $x+iy$, give a first approximation to all roots of $f(z)+c=0$.*

* The qualitative—not quantitative—aspects of many theorems in an interesting article by H. B. Mitchell, Transactions of this Society, vol. 19 (1918), pp. 43 ff., become intuitively evident by these geometric interpretations.

† See Fig. 17 for z^4-2z^2+1.

I may mention that the Fundamental Theorem of Algebra can be given a very clear interpretation by means of such space curves, and that a proof of the theorem, essentially equivalent to the proof of the existence of n intersections of $u(x, y)=0$ and $v(x, y)=0$, can be constructed.

(2) It is now necessary to free ourselves from the restriction that $a_0z^n+\cdots+a_n$ have the maximum number of real extremes. I can only indicate the state of affairs.

To illustrate, we consider successively the functions* $f(z)=z^3-z$, $z^3-\epsilon z$, z^3, $z^3+\epsilon z$, z^3+z, $0<\epsilon<1$. As $\epsilon\to 0$, the two extremes of z^3-z move together, coalescing for $\epsilon=0$, and leading to a point of inflection with a horizontal tangent ($f'(z)=0$, $f''(z)=0$).† As we pass from z^3 to z^3+z, the point of inflection is still existent, but the roots of $f'(z)=0$ have become imaginary. (For higher degree polynomials, the points of inflection generated by the coalescence of consecutive maxima and minima may be ultimately lost.) To trace this in detail requires distinction between *natural* points of inflection on polynomial curves, such as must exist between any two consecutive extremes, and *accessory* points of inflection, due to such coalescence of consecutive maxima and minima. At a *natural* point of inflection, $|f'(z)|$ is a maximum, at an *accessory* point, $|f'(z)|$ is a minimum.)‡ Schematically, $\sim$, $\backsim$ (maximum), $\backsim$, $\backsim$ (minimum).

As ϵ goes from 1 to 0 to -1, the handles first run together, and then completely separate from the real curve. The projections of our space curves onto the xy plane are given by the corresponding equations $v(x, y)=0$.

Even if the accessory points of inflection should become lost (compare, for example, the succession of curves: z^4-2z^2+1, $z^4-\epsilon z^2+1$, z^4+1, $z^4+\epsilon z^2+1$, z^4+2z^2+1) these branches are still present in the space curve. They can in no way be destroyed provided $a_0\neq 0$, but can only be changed in shape and position and in the manner in which they are connected with each other, however the coefficients a_ν may vary.

---

* See Figs. 12–16 for $\epsilon=1, 0.1, 0, -0.1, -1$, respectively.

† Our space curve consists of three branches, making with each other angles of $\pi/3$ at their point of intersection, $z=0$, as is known from the theory of conformal mapping.

‡ Kempner, Tôhoku Mathematical Journal, loc. cit. (*On the shape of polynomial curves.*)

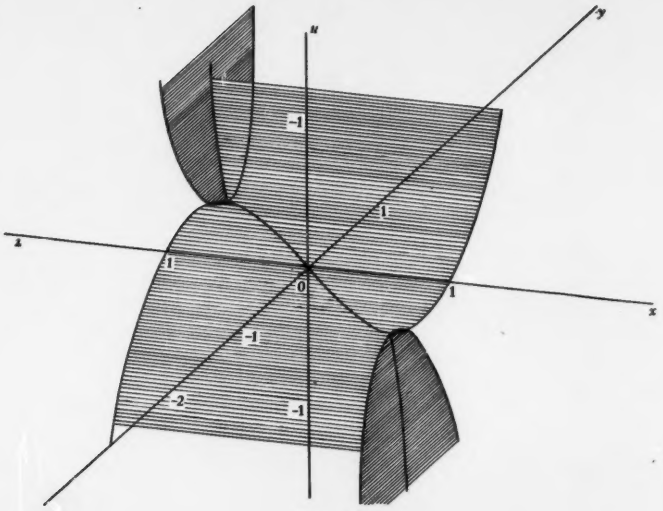

FIG. 12

FIG. 13

FIG. 14

FIG. 15

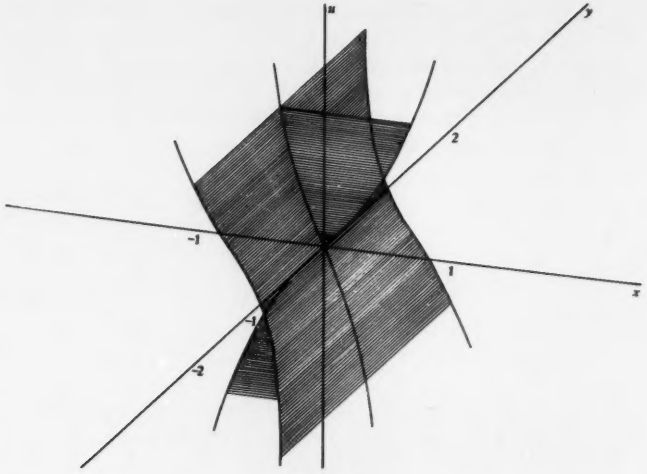

FIG. 16

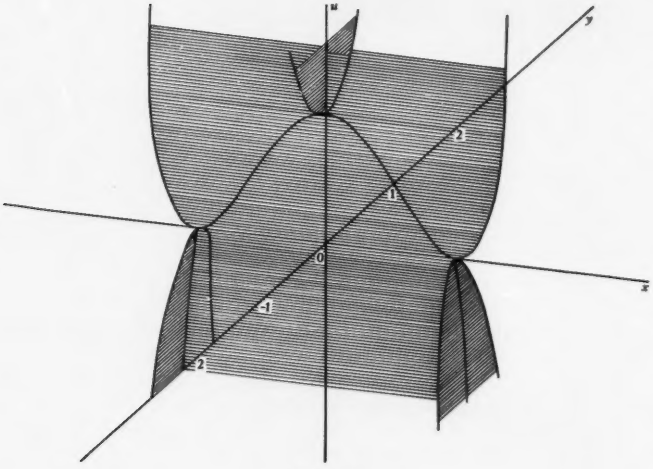

FIG. 17

These hints are sufficient to characterize qualitatively the space curves for $a_0z^n+\cdots+a_n$, for the case when the *total number* of maxima, minima, and accessory points of inflection (each accessory point given a weight 2) is the largest possible.* Algebraically, this *total number* is obtained by a careful examination of $f'(z)$ and $f''(z)$ for real z.

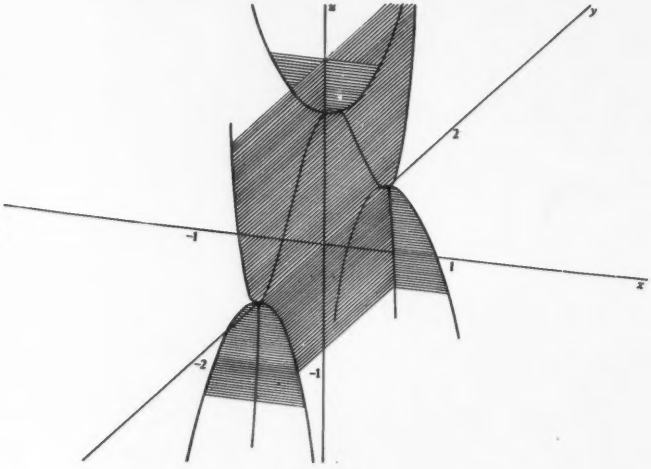

FIG. 18

(3) If our *total number* is less than $n-1$ (taking multiplicities into account), as in z^4+2z^2+1, (one minimum, no points of inflection, and hence in particular no accessory points of inflection, total number $=1<3$), some of these free branches will in turn have coalesced. As an illustration of this state of affairs, we consider z^4+2z^2+1. Since this function is obtained from z^4-2z^2+1 by a substitution $z=iz$, our whole space curve is obtained by rotating Fig. 17 through $\pi/2$ around the u axis, thus

* For a polynomial of degree n, remembering that each accessory point corresponds to one maximum and one minimum, we have $M_a+M_i+2A\leqq n-1$, where $M_a=$number of real maxima, $M_i=$number of real minima, $A=$number of accessory points. The sign of inequality holds when the curve does not possess the maximum number $n-1$ of extremes and some of the accessory points of inflection have become lost.

changing the hyperbola in the xy plane to one intersecting the y axis.

If a continuous transformation of z^4-2z^2+1 into z^4+2z^2+1 is desired which shall illustrate in typical fashion how handles attached to the real curve z^4-2z^2+1 first detach themselves from the real curve and are afterwards recaptured by it, a suitably chosen variation of coefficients passing through z^4-z^2 is preferable to our chain $z^4+\epsilon z^2$, $2 \geqq \epsilon \geqq -2$. In the general case, the situation is governed by the curve $v(x, y)=0$; any point in the xy plane where two branches of $v(x, y)=0$ intersect (double points of $v=0$) is the projection of a point on our space curve where two handles or detached branches are joined, and conversely. This agrees with the theory of conformal mapping and its breakdown at points for which $f'(z)=0$. At all such points on the space curve where handles or branches are joined (including the real curve), they are of the character of intersections of a surface at a saddle point. If a point starts moving along any branch of the space curve at $u=+\infty$, it is always possible to travel along the space curve to $u=-\infty$. (This is a consequence of the last remark combined with the Fundamental Theorem.)

I hope that enough has been said about these space curves to justify the importance I attach to them for the problem of complex roots. A collection of strong wire models of such curves, and their introduction into advanced elementary work, is very desirable.

THE UNIVERSITY OF COLORADO

A THEOREM ON HIGHER CONGRUENCES*

BY LEONARD CARLITZ

1. *Introduction.* Let $\mathfrak{D}=\mathfrak{D}(x, p^n)$ denote the totality of polynomials in an indeterminate x with coefficients in a Galois field $GF(p^n)$ of order p^n. Consider the congruence

$$t^{p^n} - t \equiv A \pmod{P}, \tag{1}$$

where A and P are in $\mathfrak{D}$, and P is irreducible of degree k, say. The sum

$$A + A^{p^n} + \cdots + A^{p^{n(k-1)}}$$

is congruent (mod P) to a quantity in $GF(p^n)$; we denote this residue by $\rho(A)$. It is easily seen that the congruence (1) is solvable in $\mathfrak{D}$ if and only if $\rho(A)=0$. A better condition is furnished by the following theorem.

THEOREM. *If we put*

$$\begin{aligned} P &= x^k + c_1x^{k-1} + \cdots + c_k, \\ P' &= kx^{k-1} + (k-1)c_1x^{k-2} + \cdots + c_{k-1}, \end{aligned} \tag{2}$$

where c_i is in $GF(p^n)$, then the congruence (1) *is solvable in* $\mathfrak{D}$ *if and only if $P'A$ is congruent* (mod P) *to a polynomial of degree $<k-1$. More generally, if*

$$P'A \equiv b_0x^{k-1} + \cdots + b_{k-1} \pmod{P}, \qquad (b_i \text{ in } GF(p^n)),$$

then $\rho(A)=b_0$.

In this note we give a new and direct proof of this theorem.†

2. *Proof of the Theorem.* For arbitrary A (mod P) we construct the polynomial

$$f(t) \equiv (t-A)(t-A^{p^n}) \cdots (t-A^{p^{n(k-1)}}) \pmod{P},$$

* Presented to the Society, April 19, 1935, under a different title.

† See L. Carlitz, *On certain functions connected with polynomials in a Galois field*, Duke Mathematical Journal, vol. 1 (1935), p. 164.

in which the coefficient of t^{k-1} is evidently $-\rho(A)$. For our purposes it will be convenient to make use of an alternative definition of $f(t)$. Let x denote a root of $P=0$; then x defines the $GF(p^{nk})$. Then $A=A(x)$ is an element of the enlarged Galois field; $f(t)$ is evidently the unique polynomial, with leading coefficient $=1$, having the roots $A^{p^{ni}}$. Clearly all the coefficients of $f(t)$ lie in $GF(p^n)$. To calculate them we proceed as follows. Let

$$x^iA \equiv \sum_{j=0}^{k-1} a_{ij}x^j \pmod{P}, \tag{3}$$

where a_{ij}, $(i, j=0, \cdots, k-1)$, are in $GF(p^n)$. But the equations (3) evidently imply the following representation of $f(t)$ as a determinant:

$$f(t) = (-1)^k \left| a_{ij} - \delta_{ij}t \right|,$$

so that by the remark at the beginning of this section

$$\rho(A) = \sum_{i=0}^{k-1} a_{ii}. \tag{4}$$

On the other hand, making use of (2) and (3),

$$\begin{aligned} P'A &= \sum_{i=1}^{k} ic_{k-i}x^{i-1}A, && (c_0 = 1), \\ &\equiv \sum_{i=1}^{k} ic_{k-i}\sum_{j=0}^{k-1} a_{i-1,j}x^j \pmod{P}, \end{aligned}$$

so that the coefficient of x^{k-1} is

$$b_0 = \sum_{i=1}^{k} ic_{k-1}a_{i-1,k-1}. \tag{5}$$

Note next that (3) implies

$$\begin{aligned} x^{i+1}A &\equiv \sum_{j=0}^{k-1} a_{ij}x^{j+1} \pmod{P} \\ &\equiv \sum_{j=0}^{k-2} a_{ij}x^{j+1} - \sum_{j=0}^{k-1} a_{i,k-1}c_{k-j}x^j, \end{aligned}$$

from which it follows that

$$a_{i+1,j} = a_{i,j-1} - a_{i,k-1}c_{k-j}. \tag{6}$$

Put $i=j-1$, and (6) becomes

$$a_{j-1,j-1} - a_{jj} = a_{j-1,k-1}c_{k-j}, \qquad (j = 1, \cdots, k-1).$$

Substituting into the right member of (5), we see that

$$\begin{aligned} b_0 &= \sum_{j=1}^{k-1} j(a_{j-1,j-1} - a_{jj}) + kc_0a_{k-1,k-1} \\ &= a_{00} + a_{11} + \cdots + a_{k-1,k-1}. \end{aligned}$$

If we compare with equation (4), we have at once $\rho(A)=b_0$. This completes the proof of the generalized form of the theorem. In particular, if $P'A$ is congruent (mod P) to a polynomial of degree $<k-1$, then $b_0=0$, and the congruence (1) is solvable.

3. *Concluding Remark.* The coefficients of $f(t)$ are, but for sign, the elementary symmetric functions of the quantities $A^{p^{nj}}$ (mod P). As we have seen above, the coefficient of t^{k-1} is intimately connected with the congruence (1). Similarly, the last coefficient

$$A^{1+p^n+\cdots+p^{n(k-1)}} \equiv \left\{\frac{A}{P}\right\} \qquad (\text{mod } P)$$

is connected with the congruence

$$t^{p^n-1} \equiv A \qquad (\text{mod } P).$$

Indeed, the method of §2 leads very naturally to F. K. Schmidt's proof of the theorem of reciprocity:

$$\left\{\frac{P}{Q}\right\} = (-1)^{kl}\left\{\frac{Q}{P}\right\},$$

where P and Q are primary irreducible of degree k and l, respectively.*

The question arises whether the remaining coefficients in $f(t)$ are connected in any direct manner with criteria for the solvability of higher congruences.

DUKE UNIVERSITY

* F. K. Schmidt, Sitzungsberichte der Physikalischmedizinischen Societät zu Erlangen, vol. 58–59 (1928), pp. 159–172.

THE HILBERT SPACE POSTULATES— A FURTHER REDUCTION

BY A. E. TAYLOR

In a recent issue of this Bulletin the writer gave a set of postulates for abstract Hilbert space in which equality was treated postulationally. The result there obtained was a set of eleven postulates, independent among themselves, defining a linear vector space of a special sort. Three more postulates were then added, to define a Hilbert space.* As a result of investigations carried on by I. E. Highberg and myself concerning postulate systems for general vector spaces,† it is now possible to make a further simplification of the Hilbert space postulates. The equality relation as an undefined idea of the system is not independent, and may be defined in terms of the other notions in the system. When this is done the number of postulates is reduced by three. The postulate system for Hilbert space is then the following, where A denotes either the complex number system C, or the real number system R.

1. *The class K contains at least one element.*
2. *If $x, y \in K$, then $x+y \in K$.*
3. *If $a \in A$ and $x \in K$, then $a \cdot x \in K$.*
4. *If $x, y \in K$, then $(x, y) \in A$.*
5. *If $x, y, z \in K$, then $(x+y, z)=(x, z)+(y, z)$.*
6. *If $x, y \in K$, then $(x, y)=\overline{(y, x)}$.*
7. *If $a \in A$ and $x, y \in K$, then $(a \cdot x, y)=a(x, y)$.*
8. *If $x \in K$, then $(x, x) \geqq 0$.*
9. *For each integer $n>0$ there exist elements $x_1, \cdots, x_n \in K$ such that $(a_1 \cdot x_1+a_2 \cdot x_2+ \cdots +a_n \cdot x_n, a_1 \cdot x_1+ \cdots +a_n \cdot x_n)=0$ if and only if $a_1=a_2= \cdots =a_n=0$.*
10. *K is separable according to the norm $(x, x)^{1/2} \equiv \|x\|$.*
11. *K is complete according to the norm $(x, x)^{1/2} \equiv \|x\|$.*

DEFINITION. *If $x, y \in K$, we say that x is equal to y and write $x=y$ if and only if $(x+-1 \cdot y, x+-1 \cdot y)=0$.*

---

* A. E. Taylor, this Bulletin, vol. 41 (1935), pp. 439–448.

† A. E. Taylor and I. E. Highberg, *On postulate systems for normed vector spaces.*

In order to prove that these postulates do define a Hilbert space we must establish the following propositions, which were postulates in the original set.

I. *If* $x, y \in K$, *and if for each* $u \in K$, $(x + -1 \cdot y) + u = u$, *then* $x = y$.

II. *If* $x \in K$ *and* $(x, x) = 0$, *then* $x + y = y$ *for each* $y \in K$.

III. *If* $x, y, u \in K$ *and* $x = y$, *then* $(x, u) = (y, u)$.

PROOF OF I. Let

$$(((x + -1 \cdot y) + u) + -1 \cdot u, ((x + -1 \cdot y) + u) + -1 \cdot u) = 0.$$

Then, since u is arbitrary, we may replace it by $0 \cdot u$. On expanding, with the aid of postulates 5, 6, 7, we easily find that $(x + -1 \cdot y, x + -1 \cdot y) = 0$, or $x = y$.

PROOF OF II. Let $(x, x) = 0$. Then, just as in the previous proof, we obtain the equation

$$((x + y) + -1 \cdot y, (x + y) + -1 \cdot y) = 0.$$

This means, by definition, $x + y = y$.

PROOF OF III. Let $(x + -1 \cdot y, x + -1 \cdot y) = 0$, and let u be an arbitrary element of K. Then $(x, u) - (y, u) = (x + -1 \cdot y, u)$. But by the Schwarz inequality, which does not depend for its proof on the use of proposition III, we have

$$0 \leqq |(x + -1 \cdot y, u)| \leqq (x + -1 \cdot y, x + -1 \cdot y)^{1/2}(u, u)^{1/2} = 0.$$

Therefore $(x, u) = (y, u)$.

Postulates 1-8 are independent among themselves, as is proved by the examples given in the previous paper. The independence of the set taken as a whole is an open question.

CALIFORNIA INSTITUTE OF TECHNOLOGY

A NOTE ON LIPSCHITZ CLASSES

BY E. S. QUADE

This note consists in the application of some results of Hardy and Littlewood* on fractional integrals to a theorem of Paley and Zygmund† and gives a generalization of that theorem.

We consider only functions of the Fourier power series type. That is, $f(x)$ is periodic in 2π, integrable, and with a Fourier series of the form

$$f(x) \sim \sum_{n=0}^{\infty} c_n e^{inx}, \qquad c_n = \frac{1}{2\pi}\int_{-\pi}^{\pi} f(x)e^{-inx}dx.$$

In dealing with functions of the class $\mathrm{Lip}(\alpha)$ or $\mathrm{Lip}(\alpha, p)$, $\alpha \neq 1$, this restriction is a matter of convenience rather than one of necessity.‡

A function $f(x)$ is said to belong to the class $\mathrm{Lip}(\alpha)$, where $0 \leqq \alpha \leqq 1$, in the interval $(-\pi, \pi)$, if

$$f(x + h) - f(x - h) = O(h^{\alpha})$$

uniformly for $-\pi \leqq x-h < x+h \leqq \pi$, and to $\mathrm{Lip}(\alpha, p)$, where $p \geqq 1$, $0 \leqq \alpha \leqq 1$, in $(-\pi, \pi)$, if $f(x) \epsilon L$, and

$$\int_{-\pi}^{\pi} |f(x + h) - f(x - h)|^p dx = O(h^{\alpha p}).$$

The functions $\phi_n(t)$, $(n = 0, 1, 2, \cdots)$, are the Rademacher functions.§

---

* Hardy and Littlewood, *Some properties of fractional integrals* I, Mathematische Zeitschrift, vol. 27 (1927–28), pp. 565–606. We will refer to this paper as (HL).

† Paley and Zygmund, *On some series of functions*, Proceedings Cambridge Philosophical Society, vol. 26 (1930), pp. 337–357. A. Zygmund, *Trigonometrical Series*, 1935, §5.61. We will refer to this book as (Z). It contains extensive bibliographical references.

‡ Hardy and Littlewood, *A convergence criterion for Fourier series*, Mathematische Zeitschrift, vol. 28 (1928), pp. 612–634, in particular, §2 and §3.5. See also (Z), §7.4.

§ For definitions and properties see (Z), §§1.32 and 5.5 to 5.61.

THEOREM 1. *Let $c_0, c_1, c_2, \cdots, c_n, \cdots$ be a sequence of real or complex numbers such that*

$$\sum_{n=2}^{\infty} n^{2\alpha} |c_n|^2 (\log n)^{1+\epsilon}$$

converges for an $\epsilon > 0$. Then, for almost all values of t, the function

$$f_t(x) \backsim \sum_{n=0}^{\infty} c_n e^{inx} \phi_n(t)$$

belongs to the class Lip(α), $(0 \leqq \alpha \leqq 1)$. *The theorem is false for the case $\alpha = 0$ or $\alpha = 1$ if $\epsilon = 0$.*

As a consequence of the theorem of Paley and Zygmund mentioned above it follows that

$$f_t^{\alpha}(x) = i^{\alpha} \sum_{n=1}^{\infty} n^{\alpha} c_n e^{inx} \phi_n(t),$$

for almost all values of t, is a continuous function (since the series converges uniformly we put $f_t^{\alpha}(x)$ equal to the sum of the series). That is, we have

$$f_t^{\alpha}(x) \epsilon \text{ Lip } (0).$$

If by the symbol $f_{t,\alpha}^{\alpha}(x)$ we denote the integral of $f_t^{\alpha}(x)$ of order α, we have*

$$f_{t,\alpha}^{\alpha}(x) \epsilon \text{ Lip } (\alpha).$$

But

$$f_{t,\alpha}^{\alpha}(x) = i^{\alpha} \sum_{n=1}^{\infty} \frac{c_n n^{\alpha}}{(in)^{\alpha}} e^{inx} \phi_n(t)$$
$$= f_t(x) - c_0 \phi_0(t).$$

To show that the theorem is not true in the case $\alpha = 1$, for $\epsilon = 0$, we consider the function

---

* (Z), §§9.80 and 9.81. A function satisfies a condition Lip*(α) or Lip*(α, p) when it satisfies a condition analogous to that for Lip(α) or Lip(α, p) but with o small in place of O large. In each of our theorems Lip(α) or Lip(α, p) may be replaced by Lip*(α) or Lip*(α, p), respectively, except in the case $\alpha = 1$; this follows from Theorems 18, 21, and 22 of (HL).

$$\sum_{m=1}^{\infty} \frac{\pm\, ie^{i2^m x}}{2^m m \log (m+1)}.$$

This can not belong to Lip(1) for any sequence of signs since it is the integral of

$$\sum_{m=1}^{\infty} \frac{\pm\, e^{i2^m x}}{m \log (m+1)},$$

which is Paley and Zygmund's example* of a series which does not represent a bounded function for any sequence of signs.

For the case of Lip (α, p), $(0 \leqq \alpha \leqq 1, p \geqq 1)$, we have a similar theorem.

THEOREM 2. *Let $c_0, c_1, c_2, \cdots, c_n, \cdots$ be a sequence of real or complex numbers such that $\sum_{n=1}^{\infty} n^{2\alpha} |c_n|^2$ converges. Then, for almost all values of t, the function*

$$f_t(x) \backsim \sum_{n=0}^{\infty} c_n e^{inx} \phi_n(t)$$

belongs to the class Lip (α, p), $(p \geqq 1, 0 \leqq \alpha \leqq 1)$.

Since $\sum_{n=1}^{\infty} |n^{\alpha} c_n|^2$ is convergent, it follows that†

$$f_t^{\alpha}(x) \backsim i^{\alpha} \sum_{n=1}^{\infty} n^{\alpha} c_n e^{inx} \phi_n(t)$$

belongs to L_p, $(p \geqq 1)$.

Now by a theorem of Hardy and Littlewood‡

$$f_{t,\alpha}^{\alpha}(x) \epsilon \text{ Lip } (\alpha, p).$$

But, as before,

$$f_{t,\alpha}^{\alpha}(x) = f_t(x) - c_0 \phi_0(t).$$

As a corollary of the following theorem we have a better theorem for the case $1 \leqq p \leqq 2$.

---

* Paley and Zygmund, loc. cit., p. 350. This gives the case $\alpha = 0$, $\epsilon = 0$.

† (Z), §5.6 (iii).

‡ (HL), Theorems 21, 22, and ff.

THEOREM 3. *If*

$$\sum_{n=-\infty}^{+\infty} n^{p'\alpha} |c_n|^{p'}, \qquad (1 < p' \leqq 2),$$

converges, then

$$f(x) \backsim \sum_{n=-\infty}^{+\infty} c_n e^{inx}$$

belongs to the class Lip (α, p), $1/p+1/p'=1$, $(0 \leqq \alpha \leqq 1)$.

From the Young-Hausdorff* theorem we have

$$\left(\frac{1}{2\pi}\int_{-\pi}^{\pi} |f^{\alpha}(x)|^{p} dx\right)^{1/p} \leqq \left(\sum_{n=-\infty}^{+\infty} |n^{\alpha} c_n|^{p'}\right)^{1/p'},$$

where

$$f^{\alpha}(x) \backsim i^{\alpha} \sum_{n=-\infty}^{+\infty} n^{\alpha} c_n e^{inx}.$$

Since $f^{\alpha}(x) \epsilon L_p$, we have†

$$f(x) - c_0 = f_{\alpha}^{\alpha}(x) = \sum_{n=-\infty}^{+\infty}{}' c_n e^{inx} \epsilon \text{ Lip } (\alpha, p), \quad (p \geqq 2).$$

COROLLARY. *If*

$$\sum_{n=-\infty}^{+\infty} n^{2\alpha} |c_n|^2, \qquad (0 \leqq \alpha \leqq 1),$$

converges, then

$$f(x) \backsim \sum_{n=-\infty}^{+\infty} c_n e^{inx}$$

belongs to Lip (α, p) *for every* p *such that* $1 \leqq p \leqq 2$.

This follows because

$$\left(\frac{1}{2\pi}\int_{-\pi}^{\pi} |f(x)|^{p} dx\right)^{1/p} \leqq \left(\frac{1}{2\pi}\int_{-\pi}^{\pi} |f(x)|^{2} dx\right)^{1/2}.$$

BROWN UNIVERSITY

* (Z), §9.1.

† (HL), Theorem 21.

A NEW PROOF OF THE EQUIVALENCE OF E. H. MOORE*

BY J. P. BALLANTINE

The theorem which is here called the equivalence of E. H. Moore is a statement of a necessary and sufficient condition for the existence of a solution of a system of linear equations.

In Moore's own proof of this equivalence, there is a hypothesis to the effect that if the Hermitian square of a vector α' is zero, all the elements are zero. That is,

$$\alpha'\bar{\alpha}_{,} = 0 \cdot \supset \cdot \alpha' = 0'.$$

In plain language, take a set of complex numbers, or numbers in any number system, multiply each by its conjugate, and add; then if the sum is zero, all the original numbers are zero.

This hypothesis is satisfied in many number systems, and holds in every case in which Moore used his equivalence. Unfortunately, the writer† used the equivalence in the number system of residues modulo 2. For such a number system the hypothesis does not hold, as for example in the case of the vector $\alpha' = (1, 1)$, where $\alpha'\bar{\alpha}_{,} = 1+1 \equiv 0$ (modulo 2). To form a logical basis for this work, it was necessary to have a proof, such as the following, of the equivalence which would not use the limiting hypothesis.

A general system of linear equations is written

$$\alpha_{,}'' \xi_{,,} = \eta_{,}$$

in matrix form. The prime range (′) and the second range (″) are conceptually distinct, so that the matrix $\alpha_{,}''$ may be square or not. Let n be the number of rows in the matrix, or the number of equations in the system, and m the number of columns, or unknowns.

The symbol $\exists$ is read "there exists," and $\ni$ is read "such that." The statement that a solution of the system of equations exists is written

---

* Presented to the Society, August 30, 1932.

† *A postulational introduction to the four color problem*, Six Studies in Mathematics, University of Washington Publications, 1930, vol. 2, No. 1, p. 9.

$$\exists \xi_{,,} \exists \alpha_{,}'' \xi_{,,} = \eta_{,}. \qquad (X)$$

Two vectors α' and $\beta_{,}$ are said to be *orthogonal* if

$$\alpha' \beta_{,} = 0.$$

A vector α' is orthogonal to all the columns of a matrix $\beta_{,}''$ if

$$\alpha' \beta_{,}'' = 0''.$$

It may be shown by straightforward substitution that

$$X \cdot \supset \cdot Y,$$

where Y is the proposition

$$\gamma' \alpha_{,}'' = 0'' \cdot \supset \cdot \gamma' \eta_{,} = 0.$$

That X implies Y and, conversely, Y implies X, constitutes the equivalence of E. H. Moore.* In Moore's proof of the converse, he normalizes the columns of the matrix, and for this step he needs the hypothesis referred to earlier. The following proof does not require this hypothesis.

To prove that Y implies X, assume Y true and X false. Then the proof is complete if we are led to a contradiction.

Commutative multiplication is assumed here, but the proof can easily be extended to a noncommutative number system. Unique division, except by zero, is also assumed.

A set, Γ_m, is built up of vectors $\gamma_{1,m}, \gamma_{2,m}, \cdots, \gamma_{N,m}$ which, together with all their linear combinations, will include all possible vectors orthogonal to all the columns of $\alpha_{,}''$. There appears to be no obvious way to write down the set Γ_m at one step, but it can be done by mathematical induction. Two sequences of sets of vectors are formed

$$A_0, A_1, \cdots, A_m,$$

$$\Gamma_0, \Gamma_1, \cdots, \Gamma_m.$$

A_k is the set of vectors consisting of the first k columns of $\alpha_{,}''$, and Γ_k is a set of vectors which, together with all their linear combinations, includes all the vectors orthogonal to all the vectors of A_k.

---

* E. H. Moore, *Lectures on General Analysis.* Memoirs of the American Philosophical Society, vol. 1, 1935.

Since A_0 is a null set, the vectors of Γ_0 may be written down, by inspection, as the rows of Hilbert's unit matrix $\delta'_{,}$, namely,

$$\Gamma_0 = (\gamma'_{1,0}, \gamma'_{2,0}, \cdots, \gamma'_{n,0}),$$

where

$$\gamma'_{1,0} = 1, 0, 0, \cdots, 0, \qquad \gamma'_{2,0} = 0, 1, 0, \cdots, 0, \qquad \cdots .$$

The formation of $\Gamma_{(k+1)}$ from Γ_k is as follows. All the vectors of Γ_k are assumed orthogonal to all those of A_k. Now $A_{(k+1)}$ involves one new vector $\nu_{,}$ not in A_k. To see if all the vectors of Γ_k are accidentally orthogonal also to $\nu_{,}$, we consider the following N_k numbers:

$$Q_j = \gamma'_{j,k}\nu', \qquad (j = 1, 2, \cdots, N_k),$$

where N_k is the number of vectors $\gamma_{1,k}, \cdots, \gamma_{N,k}$ in Γ_k.

CASE 1. Now if all the numbers Q_j are zero, then all the vectors of Γ_k are orthogonal to all the vectors of $A_{(k+1)}$. This will happen, of course, if $\nu_{,}$ the $(k+1)$st column of $\alpha'_{,}$ is linearly dependent on the preceding columns. It may, as far as we have yet shown, happen in other cases.

CASE 2. All the numbers Q_j are not zero, so that $Q_J \neq 0$, where J is some particular value of j. Let P denote the reciprocal of Q_J. In our number system such reciprocals are supposed to exist uniquely. As a tentative definition of $\Gamma_{(k+1)}$, consider

$$[\gamma'_{j,k+1}] \equiv [\gamma'_{j,k} - Q_j P \gamma'_{J,k}].$$

To test that all the vectors $\gamma'_{j,(k+1)}$ are orthogonal to all the vectors of $A_{(k+1)}$, note that it is obvious for the first k of the latter. In the case of $\nu_{,}$

$$\begin{aligned}\gamma'_{j,(k+1)}\nu_{,} &= \gamma'_{j,k}\nu_{,} - Q_j P \gamma'_{J,k}\nu_{,} \\ &= Q_j - Q_j P Q_J = 0,\end{aligned}$$

by the definition of P. In testing the linear independence of the vectors of $\Gamma_{(k+1)}$ as defined above, we find that one of them, $\gamma'_{J,(k+1)}$, is identically zero. That one must be deleted, and the others constitute the set $\Gamma_{(k+1)}$ as finally defined, so that the number of vectors in $\Gamma_{(k+1)}$ is one less than in Γ_k, or

$$N_{(k+1)} = N_k - 1.$$

The linear independence of the $N_k - 1$ vectors of $\Gamma_{(k+1)}$ follows directly from the linear independence of the vectors of Γ_k. Thus the rank of Γ_k is N_k, the same as the number of vectors in it. Let the rank of A_k be R_k. Then in Case 1,

$$N_{(k+1)} = N_k, \qquad R_{(k+1)} = R_k, \quad \text{or} \quad R_k + 1,$$

according as ν_r is or is not linearly dependent on the vectors of A_k. In Case 2,

$$N_{(k+1)} = N_k - 1, \qquad R_{(k+1)} = R_k + 1.$$

Therefore, in both cases,

$$N_{(k+1)} + R_{(k+1)} \geqq N_k + R_k.$$

When $k=0$, $N_0 = n =$ the number of rows in α_r'' and $R_0 = 0$; so that $N_0 + R_0 = n$, and therefore $N_k + R_k \geqq n$.

Now prolong the sequence $A_0, \cdots, A_m$, by defining $A_{(m+1)}$ as the set of all columns of α_r'' and the vector η_r. Then $\Gamma_{(m+1)}$ is formed from Γ_m in the manner described for forming $\Gamma_{(k+1)}$ from Γ_k.

To obtain the contradiction, we are assuming Y true and X false. Since Y is true, all the vectors of Γ_m are orthogonal to all the vectors of $A_{(m+1)}$, so we have Case 1, where $\Gamma_{(m+1)} = \Gamma_m$. Therefore $N_{(m+1)} = N_m$.

Since X is false, the last vector of $A_{(m+1)}$ is not expressible linearly in terms of the vectors of A_m. Therefore $R_{(m+1)} = R_m + 1$, so that

$$N_{(m+1)} + R_{(m+1)} > N_m + R_m \geqq n.$$

Now the sets $A_{(m+1)}$ and $\Gamma_{(m+1)}$ are mutually orthogonal, and the vectors of those sets have only n elements. It is, therefore, impossible for the sum of their ranks to exceed n, and we have the desired contradiction.

Or, by extending the sequence of A's further, one could obtain as a contradiction a single set of vectors in n-space with rank $n+1$. This, of course, could be reduced step by step to a set of vectors in one dimension of rank two.

THE UNIVERSITY OF WASHINGTON

ON THE BERNOULLI DISTRIBUTION*

BY SOLOMON KULLBACK

1. *Introduction.* The Bernoulli or binomial distribution, of central importance in the theory of mathematical probability and statistics, has been the subject of considerable study. The derivation of the moments, or recursion formulas for the moments, of this distribution, as usually presented, involve the use of moment-generating or characteristic functions, or the explicit form of the distribution itself.† In the following we will derive these moments in an elementary manner and extend the results to the Poisson exponential distribution, distributions of the Lexis and Poisson types, and the multinomial distribution.

2. *Preliminary Notions.* We need and use the following assumption, definitions, and theorem.

EMPIRICAL ASSUMPTION. If an event which can happen in two different ways be repeated a great number of times under the same essential conditions, the ratio of the number of times that it happens in one way, to the total number of trials, will approach a definite limit, as the latter number increases indefinitely.‡

DEFINITION. The limit described in the empirical assumption shall be called the probability that the event shall happen in the first way under those conditions.§ We shall express the fact

---

* Presented to the Society, November 30, 1934.

† See, for example, A. Fisher, *The Mathematical Theory of Probabilities*, 2d ed., p. 104 ff.; H. L. Rietz, *Mathematical Statistics*, 1927, p. 26 ff.; V. Mises, *Wahrscheinlichkeitsrechnung*, 1931, pp. 131–133; Risser and Traynard, *Les Principes de la Statistique Mathématique*, 1933, pp. 39–40 and 320–321; V. Romanovsky, *Note on the moments of the binomial* $(q+p)^n$ *about its mean*, Biometrika, vol. 15 (1923), pp. 410–412; A. T. Craig, *Note on the moments of a Bernoulli distribution*, this Bulletin, vol. 40 (1934), pp. 262–264; A. R. Crathorne, *Moments de la binomiale par rapport à l'origine*, Comptes Rendus, vol. 198 (1934), p. 1202; A. A. Krisknasuami Aygangar, *Note on the recurrence formulae for the moments of the point binomial*, Biometrika, vol. 26 (1934), pp. 262–264.

‡ J. L. Coolidge, *An Introduction to Mathematical Probability*, 1925, p. 4.

§ J. L. Coolidge, op. cit., p. 4.

that the relative frequency f/s (the ratio of the number of times the event happens in one way to the total number of trials) is assumed to have the probability p as a limit when $s \to \infty$, in abbreviated form by writing $E(f/s) = p$, where $E(f/s)$ is read "expected value of f/s."*

THEOREM OF COMPOUND PROBABILITY. If a compound event consists in the conjunction of any number of independent events, the probability of the compound event is the product of the probabilities for the individual events.†

3. *Bernoulli Distribution.* Suppose that in n independent trials we observe an event to occur x times. Then, applying the notions of §2, we have

$$
\begin{aligned}
&E\left(\frac{x}{n}\right) = p, \qquad E\left(\frac{x(x-1)}{n(n-1)}\right) = p^2, \cdots, \\
&E\left(\frac{x(x-1)\cdots(x-r+1)}{n(n-1)\cdots(n-r+1)}\right) = p^r, \cdots, \\
&E\left(\frac{n-x}{n}\right) = q, \qquad E\left(\frac{(n-x)(n-x+1)}{n(n-1)}\right) = q^2, \cdots, \\
&E\left(\frac{(n-x)(n-x-1)\cdots(n-x-r+1)}{n(n-1)\cdots(n-r+1)}\right) = q^r, \cdots, \\
&E\left(\frac{x(x-1)\cdots(x-r+1)(n-x)(n-x-1)\cdots(n-x-s+1)}{n(n-1)(n-2)\cdots(n-r-s+1)}\right) \\
&\qquad\qquad = p^r q^s
\end{aligned}
\tag{1}
$$

for $r+s \leqq n$. Each fraction in (1) represents the ratio of observed r-plets to the total possible number of such r-plets. We may also write (1) as

$$
\{E(x/n)\}^r \{E((n-x)/n)\}^s = E([x]^r [n-x]^s / [n]^{r+s}), \tag{2}
$$

where $[x]^r$ is the factorial function‡ $x(x-1)(x-2)\cdots(x-r+1)$

* H. L. Rietz, op. cit., p. 9.

† J. L. Coolidge, op. cit., p. 18.

‡ Whittaker and Robinson, *The Calculus of Observations*, 1924, p. 8. See also J. F. Steffensen, *Factorial moments and discontinuous frequency functions*, Skandinavisk Aktuarietidsskrift, vol. 6 (1923), pp. 73–89.

and $\{E(x/n)\}^r$ is the rth power of the expected value of x/n, and $r+s \leqq n$.

In order to find the expected value of the rth power of the variable (or the rth moment about the origin), we avail ourselves of the fact that

$$x^r = a_{r,0}[x]^r + a_{r-1,1}[x]^{r-1} + \cdots + a_{0,r}, \tag{3}$$

where $a_{i,j}$, $(i, j=0, 1, 2, \cdots)$, is the element in the ith row and jth column of the table*

0	0	0	0	0	$0\cdots$
1	1	1	1	1	$1\cdots$
1	3	7	15	31	$63\cdots$
1	6	25	90	301	$966\cdots$
1	10	65	350	$1701\cdots$	
1	15	140	$1050\cdots$		
$\cdots$	$\cdots$	$\cdots$	$\cdots$	$\cdots$	$\cdots$

in which $a_{i,j}=ia_{i,j-1}+a_{i-1,j}$, $a_{0,j}=0$, $(j=0, 1, \cdots)$, $a_{i,0}=1$, $(i=1, 2, \cdots)$. We thus have

$$E(x^r) = a_{r,0}\frac{[n]^r}{n^r}\{E(x)\}^r + a_{r-1,1}\frac{[n]^{r-1}}{n^{r-1}}\{E(x)\}^{r-1} + \cdots + a_{1,r-1}E(x), \tag{4}$$

or

$$E(x^r) = a_{r,0}[n]^r p^r + a_{r-1,1}[n]^{r-1}p^{r-1} + \cdots + a_{1,r-1}np. \tag{5}$$

For the expected value of the rth power of the deviations of x from its expected value (or the rth moment about the expected value), we have

$$E\{(x - E(x))^r\} = \sum_{i=0}^{r} (-1)^i C_{r,i} E(x^{r-i})\{E(x)\}^i, \tag{6}$$

in which we replace $E(x^r)$ by its value as given in (5). This yields

* Whittaker and Robinson, op. cit., p. 9.

$$(7) \qquad E\{(x - E(x))^r\} = B_r p^r + B_{r-1} p^{r-1} + \cdots + B_1 p,$$

where

$$B_r = \sum_{i=0}^{r} (-1)^i C_{r,i} a_{r-i,0} n^i [n]^{r-i},$$

$$B_{r-j} = \sum_{i=0}^{r-j-1} (-1)^i C_{r,i} a_{r-j-i,j} n^i [n]^{r-j-i}.$$

4. *Poisson Exponential Distribution.* The moments of the Poisson exponential distribution are derived from those of the Bernoulli distribution by letting $n \to \infty$ with $\lim_{n\to\infty} E(x) = \lim_{n\to\infty} np = m$, where m is finite. Accordingly, we have from (4) and (7), respectively,

$$(8) \qquad E(x^r) = a_{r,0} m^r + a_{r-1,1} m^{r-1} + \cdots + a_{1,r-1} m,$$

$$(9) \qquad E\{(x - E(x))^r\} = A_r m^r + A_{r-1} m^{r-1} + \cdots + A_1 m,$$

where

$$A_r = \sum_{i=0}^{r} (-1)^i C_{r,i} a_{r-i,0}, \qquad A_{r-j} = \sum_{i=0}^{r-j-1} (-1)^i C_{r,i} a_{r-j-i,j}.$$

5. *Lexis Type Distribution.** Consider N independent sets of n independent trials each, with the probability for the occurrence of an event constant for a set, but varying from set to set. Let the varying probability be $p_1, p_2, \cdots, p_N$, and the number of occurrences observed, respectively, $x_1, x_2, \cdots, x_N$.

Then, in a manner similar to that of §3, we have

$$(10) \qquad E([x]^{(r)} / [n]^r) = \frac{1}{N} \sum_{i=1}^{N} p_i^r, \quad \text{where} \quad [x]^{(r)} = \frac{1}{N} \sum_{i=1}^{N} [x_i]^r.$$

If we set $p^{(r)} = \sum_{i=1}^{N} p_i^r / N$, we have, corresponding to (5),

$$(11) \qquad E(x^{(r)}) = a_{r,0} [n]^r p^{(r)} + a_{r-1,1} [n]^{r-1} p^{(r-1)} + \cdots + a_{1,r-1} n p^{(1)},$$

where $x^{(r)} = \sum_{i=1}^{N} x_i^r / N$, and corresponding to (7),

$$(12) \qquad E\{(x^{(1)} - E(x^{(1)}))^r\} = B_r p^{(r)} + B_{r-1} p^{(r-1)} + \cdots + B_1 p^{(1)}.$$

Since†

* H. L. Rietz, op. cit., Chap. 6.

† Chrystal, *Textbook of Algebra*, 2nd ed., vol. 2, p. 49.

$$\frac{1}{N}\sum_{i=1}^{N} p_i^r \geqq \left(\frac{1}{N}\sum_{i=1}^{N} p_i\right)^r,$$

we see, by comparing (5) and (11), that

$$E_L(x^r) \geqq E_B(x^r), \tag{13}$$

where E_L represents the expected value from a Lexis type distribution and E_B the corresponding expected value from a Bernoulli distribution with $p=\sum_{i=1}^{N} p_i/N$. Indeed $E_L\{(x+a)^r\} \geqq E_B\{(x+a)^r\}$ for $a \geqq 0$.

6. *Poisson Type Distribution.** Consider a set of n independent trials with the probability for the occurrence of an event varying from trial to trial. Let these probabilities be $p_1, p_2, \cdots, p_n$. The study of the moments for this case is really the consideration of the moments for the distribution of the *total* of a Lexis type distribution of n independent sets of one trial each. Thus, if the number of observed occurrences of an event in x trials is $x=x_1+x_2+\cdots+x_n$, where $x_i=$ zero or one, then

$$E(x) = \sum_{i=1}^{n} E(x_i) = \sum_{i=1}^{n} p_i, \tag{14}$$

$$E(x^r) = a_1^{(r)}\sum p_i + a_2^{(r)}\sum p_i p_j + a_3^{(r)}\sum p_i p_j p_k + \cdots + a_r^{(r)}\sum p_i p_j \cdots p_r, \tag{15}$$

where $a_k^{(r)}=\sum r!/(s_1!s_2!\cdots s_k!)$ for $s_1+s_2+\cdots+s_k=r$, $s_i \neq 0$, $(i=1, 2, \cdots, k)$. The right member of (15) follows from the fact that $E(x_i^r)=p_i$, as may be readily seen from (4) for $n=1$ and $r \geqq 2$.

For the moments about the expected value, we have

$$E\{(x-E(x))^r\} = \sum_{i=0}^{r} (-1)^i C_{r,i} E(x^{r-i})\{E(x)\}^i, \tag{16}$$

where $E(x^r)$ is given by (15) and $\{E(x)\}^r=(p_1+p_2+\cdots+p_n)^r$.

If we repeat the preceding argument with constant probability $p=\sum_{i=1}^{n} p_i/n$, we find that

$$E(x^r) = a_1^{(r)}np + a_2^{(r)}C_{n,2}p^2 + \cdots + a_r^{(r)}C_{n,r}p^r. \tag{17}$$

* H. L. Rietz, op. cit., Chap. 6.

Since*

$$C_{n,r}\left(\frac{1}{n}\sum_{i=1}^{n}p_i\right)^r \geqq \sum p_i p_j \cdots p_r,$$

we see by comparing (17) and (15) that

$$E_B(x^r) \geqq E_P(x^r), \tag{18}$$

where E_P represents the expected value from a Poisson type distribution and E_B the corresponding expected value from a Bernoulli distribution with $p=\sum_{i=1}^{n}p_i/n$. Indeed, $E_B\{(x+a)^r\} \geqq E_P\{(x+a)^r\}$ for $a \geqq 0$.

A comparison of (17) and (5) shows that

$$C_{n,k}a_k^{(r)} = [n]^k a_{k,r-k}, \quad \text{or} \quad a_k^{(r)} = k!a_{k,r-k}. \tag{19}$$

7. *The Multinomial Distribution.* Consider a trial in which one of r mutually exclusive events may occur. Let the respective probabilities of occurrence be $p_1, p_2, \cdots, p_r$ with $p_1+p_2+\cdots+p_r=1$. If in n independent trials there are observed $x_1, x_2, \cdots, x_r$ occurrences, respectively, of the r events, then

$$E\{[x_1]^\alpha[x_2]^\beta \cdots [x_r]^\rho / [n]^{\alpha+\beta+\cdots+\rho}\} = p_1^\alpha p_2^\beta \cdots p_r^\rho. \tag{20}$$

Thus, for example, $E(x_ix_j)=n(n-1)p_ip_j$ and $E(x_ix_j)-E(x_i)E(x_j) = n(n-1)p_ip_j-np_inp_j=-np_ip_j$.

Now we may write†

$$x_i^l = \sum_{\lambda=0}^{l}\Delta^\lambda 0^l[x_i]^\lambda/\lambda!,$$

where $\Delta^\lambda 0^l$ are the differences of zero and

$$\Delta^\lambda 0^l = \lambda^l - \lambda(\lambda-1)^l + \frac{1}{2!}\lambda(\lambda-1)(\lambda-2)^l - \cdots \pm \lambda\cdot 1^l.$$

---

* This can be shown by using the fact that $f(x_1, x_2, \cdots, x_n)=\sum x_1x_2\cdots x_r$ subject to the condition $x_1+x_2+\cdots+x_n=$constant, has its maximum value for $x_1=x_2=\cdots=x_n$.

† J. F. Steffensen, *Interpolation*, 1927, p. 13, equation (3).

Accordingly,

$$(21)\qquad x_1^{k}x_2^{l}\cdots x_r^{m} = \sum_{\kappa=0}^{l}\frac{\Delta^{\kappa}0^{k}[x_1]^{\kappa}}{\kappa!}\sum_{\lambda=0}^{l}\frac{\Delta^{\lambda}0^{l}[x_2]^{\lambda}}{\lambda!}\cdots\sum_{\mu=0}^{m}\frac{\Delta^{\mu}0^{m}[x_r]^{\mu}}{\mu!},$$

and

$$(22)\qquad E(x_1^{k}x_2^{l}\cdots x_r^{m}) = \sum_{\kappa=0,\lambda=0,\cdots,\mu=0}^{k,l,\cdots,m}\frac{\Delta^{\kappa}0^{k}\cdot\Delta^{\lambda}0^{l}\cdots\Delta^{\mu}0^{m}[n]^{\kappa+\lambda+\cdots+\mu}p_1^{\kappa}p_2^{\lambda}\cdots p_r^{\mu}}{\kappa!\lambda!\cdots\mu!}$$
$$= (1+p_1\Delta_1+p_2\Delta_2+\cdots+p_r\Delta_r)^{n}\cdot 0_1^{k}0_2^{l}\cdots 0_r^{m},$$

where $\Delta_i^{k}0_j^{l}=1$ for $i\neq j$ and $\Delta_i^{k}0_i^{l}=\Delta^{k}0^{l}$. The latter result is true for any n and $k, l, \cdots, m$ since $[n]^{\kappa+\lambda+\cdots+\mu}=0$ for $\kappa+\lambda+\cdots+\mu>n$ and $\Delta^{\alpha}0^{r}=0$ for $\alpha>r$.

The procedure with respect to the product moments about any constant is similar to that outlined. Thus, since* $(x_i-a_i)^{l}=\sum_{\lambda=0}^{l}\Delta^{\lambda}(-a_i)^{l}[x_i]^{\lambda}/\lambda!$, we have

$$(23)\qquad (x_1-a_1)^{k}(x_2-a_2)^{l}\cdots(x_r-a_r)^{m} = \sum_{\kappa=0,\lambda=0,\cdots,\mu=0}^{k,l,\cdots,m}\frac{\Delta^{\kappa}(-a_1)^{k}\Delta^{\lambda}(-a_2)^{l}\cdots\Delta^{\mu}(-a_r)^{m}[x_1]^{\kappa}[x_2]^{\lambda}\cdots[x_r]^{\mu}}{\kappa!\lambda!\cdots\mu!},$$

and

$$(24)\qquad E\{(x_1-a_1)^{k}(x_2-a_2)^{l}\cdots(x_r-a_r)^{m}\} = (1+p_1\Delta_1+p_2\Delta_2+\cdots+p_r\Delta_r)^{n}\cdot(-a_1)^{k}(-a_2)^{l}\cdots(-a_r)^{m},$$

where $a_1, a_2, \cdots, a_r$ are any constants and $\Delta_i^{k}(-a_j)^{l}=1$ for $i\neq j$ and $\Delta_i^{k}(-a_i)^{l}=\Delta^{k}(-a_i)^{l}$. Since† $\Delta=e^{D}-1$, where $D=d/dx$,

* J. F. Steffensen, op. cit., p. 13.

† Whittaker and Robinson, op. cit., p. 63.

$$(25)\qquad E\{(x_1-a_1)^k(x_2-a_2)^l\cdots(x_r-a_r)^m\} = (p_1e^{D_1}+p_2e^{D_2}+\cdots+p_re^{D_r})^n\cdot x_1^k x_2^l\cdots x_r^m\Big]_{\substack{x_1=a_1\\ \vdots\\ x_r=a_r}},$$

where $D_1=\partial/\partial x_1$.

WASHINGTON, D. C.

ON A RESULTANT CONNECTED WITH FERMAT'S LAST THEOREM

BY EMMA LEHMER

E. Wendt* seems to have been the first to introduce the resultant of $x^n=1$ and $(x+1)^n=1$ in connection with Fermat's Last Theorem. This resultant can be expressed by means of the following circulant of binomial coefficients

$$\Delta_n=\begin{vmatrix} 1 & C_{n,1} & C_{n,2}\cdots C_{n,n-1}\\ C_{n,n-1} & 1 & C_{n,1}\cdots C_{n,n-2}\\ \cdots & \cdots & \cdots\\ C_{n,1} & C_{n,2} & C_{n,3}\cdots 1\end{vmatrix}.$$

In his book on Fermat's Last Theorem Bachmann† proved that if p is an odd prime and if Δ_{p-1} is not divisible by p^2, then Fermat's equation $x^p+y^p+z^p=0$ has no solution (x, y, z) prime to p.

S. Lubelsky‡ proved in a recent paper, using the distribution of quadratic residues, that if $p\geqq 7$, Δ_{p-1} is not only divisible by p^2, but by p^3, thus annulling Bachmann's criterion except for $p=3$ and $p=5$.

We shall now show how, by a straightforward manipulation with the above determinant, one can prove much more.

THEOREM 1. *Δ_{p-1} is divisible by $p^{p-2}q_2$ for every prime p, where q_2 is the Fermat quotient $(2^{p-1}-1)/p$.*

---

* Journal für Mathematik, vol. 113 (1894), pp. 335–347.

† *Das Fermatproblem*, 1919, p. 59.

‡ Prace Matematyczno-Fizyczne, vol. 42 (1935), pp. 11–44.

PROOF. First add to each element of the last column of Δ_{p-1} the corresponding element of all the other columns. The elements of the last column now become equal to

$$1 + C_{p-1,1} + C_{p-1,2} + \cdots + C_{p-1,p-2} = 2^{p-1} - 1.$$

Next increase each element of the first $p-3$ columns by the element immediately to the right of it. All the elements of the first $p-3$ columns are now of the form

$$C_{p-1,k} + C_{p-1,k+1} = C_{p,k+1} = pI_k, \quad (k = 0, 1, \cdots, p-2).$$

Since, as is well known, I_k is an integer for p a prime, it follows that p is a factor of each of the first $p-3$ columns. Also $(2^{p-1}-1)$ comes out of the last column, and hence Δ_{p-1} is divisible by $(2^{p-1}-1)p^{p-3} = p^{p-2}q_2$, which is the theorem.

For example, if $p=5$,

$$\Delta_4 = \begin{vmatrix} 1 & 4 & 6 & 4 \\ 4 & 1 & 4 & 6 \\ 6 & 4 & 1 & 4 \\ 4 & 6 & 4 & 1 \end{vmatrix} = \begin{vmatrix} 1 & 4 & 6 & 15 \\ 4 & 1 & 4 & 15 \\ 6 & 4 & 1 & 15 \\ 4 & 6 & 4 & 15 \end{vmatrix}$$

$$= \begin{vmatrix} 5 & 10 & 6 & 15 \\ 5 & 5 & 4 & 15 \\ 10 & 5 & 1 & 15 \\ 10 & 10 & 4 & 15 \end{vmatrix} = 15 \cdot 5^2 \begin{vmatrix} 1 & 2 & 6 & 1 \\ 1 & 1 & 4 & 1 \\ 2 & 1 & 1 & 1 \\ 2 & 2 & 4 & 1 \end{vmatrix} = -5^3 \cdot 3.$$

It is interesting to notice that although Theorem 1 is in no way dependent on the solvability of Fermat's equation, nevertheless it enables us to replace Bachmann's criterion by the following one.

If Δ_{p-1} is not divisible by p^{p-1}, then $x^p+y^p+z^p=0$ has no solution (x, y, z) prime to p.

This in fact is merely a restatement of Wiefrich's criterion,* which states that if p does not divide q_2, Fermat's equation has no solution in integers prime to p.

For example, since we have seen that Δ_4 is not divisible by 5^4, it follows from our criterion that $x^5+y^5+z^5=0$ has no solutions (x, y, z) prime to 5.

* Journal für Mathematik, vol. 136 (1909), pp. 293–302.

Before proceeding further it may be of interest to give a short, though perhaps less elementary proof of Theorem 1, based directly on the definition of the resultant of two polynomials as the product $f(\alpha_1)\cdot f(\alpha_2)\cdots f(\alpha_n)$, where f is one polynomial and the α's are roots of the other.

$$\Delta_{p-1} = \prod_{j=1}^{p-1}[(\epsilon_j + 1)^{p-1} - 1],$$

where ϵ_j are all the $(p-1)$st roots of unity.

For $\epsilon_j = 1$, we get the factor $2^{p-1}-1$, and for $\epsilon_j = -1$, we get -1. For $\epsilon_j = \epsilon$, any complex root of unity, we have

$$(\epsilon + 1)^{p-1} - 1 = C_{p-1,1}\epsilon + C_{p-1,2}\epsilon^2 + \cdots + \epsilon^{p-1}.$$

But since, for p a prime,

$$C_{p-1,k} = (-1)^k + pc_k,$$

where the c's are integers,* we have

$$\begin{aligned}(\epsilon + 1)^{p-1} - 1 &= -\epsilon + \epsilon^2 - \epsilon^3 + \cdots + \epsilon^{p-1} + pf(\epsilon)\\ &= \epsilon(\epsilon^{p-1} - 1)/(\epsilon + 1) + pf(\epsilon) = pf(\epsilon),\end{aligned}$$

where $f(x)$ is a polynomial with integral coefficients. Hence

$$\Delta_{p-1} = -(2^{p-1} - 1)\prod_{\epsilon_j \neq \pm 1} pf(\epsilon_j) = -(2^{p-1} - 1)\cdot p^{p-3}\prod f(\epsilon_j),$$

where $\prod f(\epsilon_j)$ is an integral symmetric function of the roots of $(x^{p-1}-1)/(x^2-1)$ and hence an integer. Hence Δ_{p-1} is divisible by $p^{p-2}q_2$.

In comparing Theorem 1 with that of Lubelsky we see that Theorem 1 says more, except for $p=7$, in which case Theorem 1 guarantees divisibility of Δ_6 by 7^5 instead of 7^8. On examining this case more closely we find that as a matter of fact $\Delta_6 = 0$. Indeed, we have in general the following theorem.

THEOREM 2. $\Delta_n = 0$ *if and only if* $n = 6k$.

PROOF. In order that $\Delta_n = 0$ it is both necessary and sufficient that $x^n = 1$ and $(x+1)^n = 1$ have a root ρ in common. But the roots of $x^n = 1$ are the nth roots of unity. Hence we can write

* Lucas, American Journal of Mathematics, vol. 1 (1878), pp. 229–230.

$\rho=\cos\theta+i\sin\theta$, but since at the same time ρ is a root of $(x+1)^n=1$, we have

$$|\rho+1| = 1 = (\cos\theta+1)^2+\sin^2\theta = 2+2\cos\theta,$$

or $\cos\theta=-1/2$ and $\theta=\pm 2\pi/3$. This condition will be satisfied if and only if $\rho=\omega$ or ω^2, while $(\rho+1)=-\omega^2$ or $-\omega$, and hence $(\rho+1)^n=1$ if and only if n is a multiple of 6.

One can easily show by adding to each element of the first and second column of Δ_{6k} (written in determinant form) the corresponding elements of every third column, that the elements of the two resulting columns will be equal and hence that $\Delta_{6k}=0$, but the writer has not been able to show from the circulant definition of Δ_n that $\Delta_n\neq 0$ if n is not a multiple of 6.

In conclusion we give a short table of Δ_{p-1}.

p	Δ_{p-1}
3	$-3=-(2^2-1)$
5	$-375=-(2^4-1)\cdot 5^2$
7	0
11	$-210\ 736\ 858\ 987\ 743=-(2^{10}-1)\cdot 11^8\cdot 31^2$
13	0
17	$-1\ 562\ 716\ 604\ 038\ 367\ 719\ 196\ 682\ 456\ 673\ 375=-(2^{16}-1)\cdot 17^{14}\cdot(3^3\cdot 5\cdot 7^3\cdot 257)^2$
19	0

It appears from this table of Δ_{p-1}, and can be shown without difficulty for any even n, that Δ_n is $-(2^n-1)$ times a perfect square. It can also be shown that Δ_d divides Δ_n if d divides n.

BETHLEHEM, PENNSYLVANIA

CORRESPONDENCES CONNECTED WITH A PENCIL OF n-ICS*

BY A. R. WILLIAMS

1. *Introduction.* All curves of order n passing through $(n/2)(n+3)-1$ points pass through $(n-1)(n-2)/2$ other points. That is, the n^2 base points of a pencil of n-ics are determined by the former number. When $(n/2)(n+3)-2$ are fixed in general position, and another moves on a curve of order m, the locus of the remaining $(n-1)(n-2)/2$ is a curve of order $m(n^2-1)$, which has a multiple point of order mn at each of the fixed points.† In a previous paper I have considered the case when a number of base points are fixed and the others necessary to determine a pencil are taken consecutive on some curve.‡ One other situation is perhaps worth brief notice. Of the nq intersections of two curves of order n and q, $n>q$, $nq-(q-1)(q-2)/2$ can be taken arbitrarily, the others being then determined. Moreover,

$$nq - \frac{(q-1)(q-2)}{2} + \frac{(n-q)(n-q+3)}{2} - 1 = \frac{n(n+3)}{2} - 2.$$

This means that if $nq-(q-1)(q-2)/2$ base points are taken on a curve of order q, and $(1/2)(n-q)(n-q+3)-1$ others not on the latter, then if one other, say P, describes some locus, $(q-1)(q-2)/2$ will be fixed on the curve of order q, and the remaining $(1/2)(n-1)(n-2)-(1/2)(q-1)(q-2)$ $=(1/2)(n-q)(n+q-3)$, variable with P, will, for any position of P, lie on the curve of order $n-q$ determined by P and the second set of fixed points just mentioned. It is the purpose of this paper to discuss the locus of the variable points when P describes a given curve.

2. *Order of the Locus and Its Singularities at the Base Points.* In this connection we may use the rational surface whose plane

---

* Presented to the Society, April 6, 1935.

† Milinowski, Journal für Mathematik, vol. 77, p. 263.

‡ Williams, this Bulletin, vol. 36, p. 133.

sections correspond to the ∞^3 n-ics that have in common the nq fixed points on the q-ic and all but one of the others. It will be convenient to refer to the set on the q-ic as the points A, and the others, $(1/2)(n-q)(n-q+3)-2$ in number, as the points B. The remaining fixed point, which reduces the multiplicity of n-ics to a net, may be called Q. The points B and Q determine a pencil of $(n-q)$-ics. The order of the surface is $(1/2)(n-q)(n+q-3)+2$. It has a multiple point of order $q(n-q)$ corresponding to the q-ic. The sections through the multiple point correspond to the net of $(n-q)$-ics determined by the points B. In particular, to the points A on the q-ic correspond nq lines of the surface which pass through the multiple point. To a pencil of n-ics determined by the points A, B, Q, and P corresponds a pencil of plane sections whose axis is $Q'P'$, the latter being the points of the surface that correspond to Q and P. The axis $Q'P'$ meets the surface in $(1/2)(n-q)(n+q-3)$ other points to which correspond in the plane the remaining base points of the pencil of n-ics. As P moves on a general line l, P' moves on an n-ic, which is skew unless $q=n-1$, and has a multiple point of order q at the multiple point of the surface. Thus the cone generated by $Q'P'$ has a q-fold edge. It meets the surface in the skew n-ic and a curve whose image in the plane is the locus sought. To find the order of the latter we have only to note that to a second line l' of the plane corresponds a similar n-ic on the surface which meets the cone described by $Q'P'$ in n^2 points, q^2 of which are at the multiple point of the surface, while another corresponds to the intersection of l and l'. Hence as P moves on l, the variable points fill a curve L of order n^2-q^2-1 which passes $n-q$ times through each of the points A, n times through each of the points B, and n times through the point Q, each passage through Q corresponding to an intersection of the locus of P' with the tangent plane to the surface at Q'. Thus, of the intersections of L with the q-ic, $nq(n-q)$ are accounted for at the points A; and the remaining $q[q(n-q)-1]$ correspond to the q intersections of l with the q-ic. When P takes a position on the q-ic, the pencil of n-ics becomes the q-ic and the pencil of $(n-q)$-ics determined by the points B and Q. One of these passes through P and meets the q-ic in $q(n-q)-1$ other points which correspond to P. The others are the $(1/2)(n-q-1)(n-q-2)$ remaining base points of the

pencil of $(n-q)$-ics. The sum of the last two expressions is $(1/2)(n-q)(n+q-3)$, the number of points variable with P. The remaining base points of the pencil of $(n-q)$-ics correspond to the points in which the line joining Q' to the node of the surface meets the latter again. Each of them is a q-fold point on L, since P in describing l meets the q-ic q times.

3. *Self-Corresponding Points and Branch-Points.* The locus of of points P such that one of the corresponding points coincides with P is the jacobian of the net of n-ics determined by the points A, B, and Q. It is the image of the curve of contact of the tangent cone drawn to the surface from Q'. This jacobian contains the q-ic. The residual J, of order $3(n-1)-q$, meets l in those points which are points of L. Of the other intersections of J and L, $n-q$ are accounted for at each of the nq points A, and $2n$ at each of the points B and at Q. This leaves $(n-q)(2n^2+2nq-6n-3q-q^2-2)-2(n-3)$ points of intersection which do not take place on l or at any of the base points. These are points at which two of the points corresponding to a position of P coincide. That is, they correspond to *branch-points* on l. In connection with the correspondence between the points of l and L we may apply the formula of Zeuthen, $\eta-\eta'=2\alpha(p'-1)-2\alpha'(p-1)$. Here α and α' are 1 and $(n-q)(n+q-3)/2$, while η and η' are the numbers of branch-points on l and L, and p and p' are their genera. That is, η is the number last determined, η' is 0, and p is 0. Thus p', the genus of L, is $(\eta/2)-\alpha'+1$, which is found to be $(1/2)(n-q)(2n^2+2nq-q^2-4q-7n+1)-(n-4)$. It is of some interest to verify this number. The line l, the locus of P, and the corresponding locus L meet in $n^2-q^2-1-3(n-1)+q$ or $(n-1)(n-2)-q(q-1)$ points which are not on J. These come in pairs, each pair corresponding to an apparent double point of the locus of P' as seen from Q'. Moreover, with each such pair are associated $\alpha'-1$ or $(1/2)(n-q)(n+q-3)-1$ nodes on L. For any one of the $\alpha'+1$ associated points determines the others. Let P_1 be a point of l for which one of the corresponding points, say R_1, falls on l. When P describing l reaches R_1, one corresponding point is at P_1 and the $\alpha'-1$ others coincide with points corresponding to P_1, producing $\alpha'-1$ nodes. Now the genus of L found by considering the singularities at the points A, B, and Q, the $(1/2)(n-q-1)(n-q-2)$ q-fold points common to the pencil

of $(n-q)$-ics, and the $(1/2)[(n-1)(n-2)-q(q-1)](\alpha'-1)$ nodes just mentioned, agrees with the above value of p' given by the formula of Zeuthen.

If now P moves on a curve c of order m, the locus of the corresponding points is a curve C of order $m(n^2-q^2-1)$ which has an mn-fold point at each of the points B and at Q, and an $m(n-q)$-fold point at each of the points A. Each of the remaining base points of the pencil of $(n-q)$-ics is an mq-fold point of C. The number of branch-points on c is also m times the number on l. Zeuthen's relation then gives for the genus of C, $p'=(\eta/2)-\alpha'+\alpha' p+1$. Here p is the genus of c, and η, the number of branch-points on c, is m times the number of branch-points on l. This shows, as is geometrically evident, that each node of c gives α' nodes on C. The curves c and C meet in $m(n^2-q^2-1)-m(3n-3-q)$ points not on J. The last number may be written $(mn-1)(mn-2)-mq(mq-1)-(m-1)(m-2)$, which shows that it is twice the number of apparent double points of the locus of P'. With each such apparent double point, or with each such pair of intersections of c and C, are associated $\alpha'-1$ nodes of C. Moreover, as just remarked, each of the $(1/2)(m-1)(m-2)-p$ nodes or cusps of c gives α' nodes or cusps on C. Then, considering the singularities of C at the points A, B, and Q, the mq-fold points at the remaining base points of the pencil of $(n-q)$-ics, and the two classes of simple nodes just described, the genus of C is found to agree with the above. It is easily shown that if the locus of P passes through a point B or through Q, the corresponding locus C loses the n-ic of the net that has that point for double point. But if the locus of P passes through a point A on the q-ic, C loses the $(n-q)$-ic that is determined by the points B, Q, and the point A in question.

4. *Particular Cases.* The cases $q=n-1$ and $q=n-2$ are of interest. When $q=n-2$, there are only three points B. The number of base points variable with P is $2n-5$. If P describes a line, the order of their locus is $4n-5$. The order of the rational surface is $2n-3$, and its multiple point is of order $2n-4$, corresponding to the fact that the only base points of the pencil of conics are the three points B and Q. The conics of the net determined by the points B are the images of sections by planes through the multiple point. The double curve of the surface

consists of three $(n-2)$-fold lines passing through the multiple point. Their images are the lines of the plane joining the points B in pairs. That is, through the points A and B pass a web of n-ics, and the $n-2$ remaining intersections of an n-ic of this web with the line joining two points B are determined by one of them. When P moves on the line B_1B_2, the locus of the variable points is B_1B_2 taken $n-3$ times and B_3Q taken $n-2$ times.

When $q=n-1$, there are no points B. When Q is fixed, the number of points variable with P is $n-2$. They lie on the line QP; and if P moves on a general line, their locus is of order $2(n-1)$, and has an n-fold point at Q and simple points at the points A on the $(n-1)$-ic. The rational surface is of order n, and has a multiple point of order $n-1$. Sections through the latter correspond to the lines of the plane. If P describes a line passing through Q, the locus of the variable points is of course that line. But in this connection one fact may be noted. To the line joining Q and A_k on the $(n-1)$-ic corresponds on the surface a plane $(n-1)$-ic whose plane contains the line a_k of the surface that corresponds to A_k. Hence, for any position of P on the line QA_k one of the associated points is at A_k. That is, the pencil of n-ics have a common tangent at A_k, the one having a double point being the $(n-1)$-ic and the line.

Since the number of points A on the $(n-1)$-ic is just sufficient to determine a web of n-ics, we may take Q, P consecutive on some curve c of order m. That question for the base points in general position was considered in the article referred to at the beginning of the paper. The modification to be made may be indicated briefly. To c corresponds on the surface a curve c', of order mn, having an $m(n-1)$-fold point at the $(n-1)$-fold point of the surface. To a pencil of plane n-ics tangent to c corresponds a pencil of plane sections whose axis is tangent to c' and meets the surface in $n-2$ other points, giving a curve on the surface whose image in the plane is the locus sought. The tangents to c' form a developable of order $2(mn-1+p)$, where p is the genus of c and c'. For the sake of simplicity c is supposed to have no cusps and to contain none of the base points. Projecting c' from the multiple point of the surface, which we have seen is of order $n-1$ for the surface and of order $m(n-1)$ for c', we find that the developable has there a multiple point of order $2m(n-1)$. The developable intersects the surface in the

cuspidal edge counted twice, and in a residual curve whose image is the locus sought. To find the order of the latter we have only to note that to a line of the plane corresponds on the surface a plane n-ic which has an $(n-1)$-fold point at the multiple point of the surface. Multiplying the order of the developable by n, and subtracting the $2m(n-1)^2$ intersections at the multiple point and the $2m$ on the cuspidal edge, we find that the order of the locus sought is $2(2mn-2m-n+np)$. This holds for $m \geqq 2$. The locus passes $2(m+p-1)$ times through each of the points A. If the points A were in general position, the order r of the developable would remain the same; but the order of the locus would be $rn-2m$, that is, greater by $2m(n-1)^2$, and it would pass $2m(n-1)$ more times through each of the points A. In addition to the singularities at the base points the locus has a number of nodes corresponding to those intersections of the double curve of the developable with the rational surface that do not occur on the cuspidal edge. Since the genus of a plane section of the developable is the same as that of the cuspidal edge, the order of the double curve is $(1/2)(r-1)(r-2)-mn-p$. If the cuspidal edge has a multiple point of order s, the double curve has there a multiple point of order $2s(s-1)$. Therefore, the double curve of the developable in question has a multiple point of order $2m(n-1)[m(n-1)-1]$ at the multiple point of the rational surface, and a multiple point of order 4 at each of the $(1/2)(m-1)(m-2)-p$ nodes of c'. The number α of stationary planes of the cuspidal edge is given by $3(r-2)-2mn+6p$. The point of contact of such a plane is a simple point on both the double curve and the cuspidal edge. A point where the cuspidal edge is met by a tangent to itself is a cusp on the double curve. A curve in 3-space of order μ and genus p without nodes or cusps has $2(\mu-2)(\mu-3)+2p(\mu-6)$ tangents that meet it again. This number is reduced by 4 when an apparent double point is replaced by a node, and by the order μ if the apparent double point is replaced by a cusp. In the present instance the order of the cuspidal edge is mn, the genus is p, and there is a multiple point of order $m(n-1)$ and $(1/2)(m-1)(m-2)-p$ nodes. The number γ of tangents to the cuspidal edge that meet it again is thus found to be $4m(1+mn-m-2n)+8+2p(mn-4)$. The number of intersections of the double curve of the developable with the ra-

tional surface that do not take place on the cuspidal edge is $nx-2m(n-1)^2[m(n-1)-1]-\alpha-2\gamma-4[(1/2)(m-1)(m-2)-p]$, where x is the order of the double curve given above. Each such intersection corresponds to a node on the locus sought. Therefore, considering this number of nodes and the singularities at the base points, the genus of the locus is found to be $(n-3)(3mn-2m-n-3)+p(n^2-8)$. If the base points were in general position, the corresponding expression would be as follows:* $(n-3)(5mn^2-6m-5n)/2+p(5n^2-15n+2)/2$. As is to be expected, both reduce to p when $n=3$. For we should then be dealing with a pencil of cubics having six fixed base points while Q, P, the seventh and eighth, are consecutive on the curve of order m. Hence to each point of this curve corresponds one point on the locus in question. And the genus of the latter we should expect to be the same whether the six fixed points are on a conic or not.

Finally if $m=1$, that is, if the two consecutive points P, Q describe a line l, the locus of the others is that same line, and the above result does not hold. But since to l there corresponds on the surface a rational plane n-ic having an $(n-1)$-fold point, it is seen immediately that the number of positions of P on l such that one of the corresponding points coincides with P is the number of inflections of such an n-ic, that is, $3(n-2)$. And the number of positions of P such that two corresponding points coincide is twice the number of double tangents, that is, $4(n-2)(n-3)$.

THE UNIVERSITY OF CALIFORNIA

* See the paper mentioned above.

SOME MULTIPLICATION THEOREMS FOR THE NÖRLUND MEAN

BY FLORENCE M. MEARS

Absolute summability for the series $\sum_{n=1}^{\infty} u_n$ by the Cesàro mean and by the Riesz mean have been defined by Fekete* and by Obrechkoff,† respectively. In each case, theorems for the multiplication of series summed by these means have been proved.‡ The purpose of this paper is to establish a definition for absolute summability by the Nörlund mean, and to prove three multiplication theorems for this mean. Theorem 1 includes Mertens' theorem for convergent series and its extension for the Cesàro mean. Theorem 2 includes Cesàro's multiplication theorem. Theorem 3 includes the following theorem by M. J. Belinfante for the Cesàro mean.

If $\sum_{n=1}^{\infty} u_n$ *is summable* C_s *to* U, *and if* $\sum_{n=1}^{\infty} v_n$ *is summable* C_r *to* V, *and bounded* C_{r-1}, $(s \geqq 0, r \geqq 1)$, *the product series* $\sum_{n=1}^{\infty} w_n$ *is summable* C_{r+s} *to* UV.§

For any given series $\sum_{k=1}^{\infty} u_k$, with terms real or complex, form the sequence $\{U_k\}$, where $U_k = \sum_{n=1}^{k} u_n$. Let $\{a_n\}$ be a sequence of positive numbers, and let $A_k = \sum_{n=1}^{k} a_n$. The series $\sum_{k=1}^{\infty} u_k$ is said to be summable to U' by the Nörlund mean A if

$$\lim_{n\to\infty} U_n' = \lim_{n\to\infty} \frac{\sum_{k=1}^{n} a_{n-k+1} U_k}{A_n}$$

exists and is equal to U'.¶ If $\sum_{k=1}^{\infty} u_k'$, where $u_n' = U_n' - U_{n-1}'$, is absolutely convergent, we shall say that $\sum_{k=1}^{\infty} u_k$ is absolutely summable A. We shall assume that $\lim_{n\to\infty}(a_n/A_n) = 0$; then A is a regular method of summation.‖

* Matematikai és Természettudományi Értesitö, vol. 32 (1914), pp. 389–425.

† Comptes Rendus, vol. 185 (1928), pp. 215–217.

‡ For discussion and references, see Kogbetliantz, Mémorial des Sciences Mathématiques, No. 51.

§ Koninklijke Akademie te Amsterdam, Verslag, vol. 32 (1923), pp. 177–189.

¶ Riesz, Proceedings of the London Mathematical Society, (2), vol. 22 (1923), p. 412.

‖ Riesz, loc. cit.

We shall consider also the series $\sum_{k=1}^{\infty} v_k$, and $\sum_{k=1}^{\infty} w_k$, the Cauchy product of series $\sum_{k=1}^{\infty} u_k$ and $\sum_{k=1}^{\infty} v_k$; we have the corresponding sequences $\{V_k\}$ and $\{W_k\}$.

We shall assume that we have also a regular Nörlund mean, B, defined by $\{b_k\}$, a sequence of positive numbers. We shall form the Nörlund means, C, defined by $\{c_n\} = \{\sum_{k=1}^{n} a_k b_{n-k+1}\}$, and D, defined by $\{d_n\} = \{\sum_{k=1}^{n} A_k b_{n-k+1}\}$.

THEOREM 1. *If $\sum_{k=1}^{\infty} u_k$ is summable A to U', and in addition, absolutely summable A, and if $\sum_{k=1}^{\infty} v_k$ is summable B to V', then $\sum_{k=1}^{\infty} w_k$ is summable C to $U'V'$.*

PROOF. We shall prove the theorem first with the assumption that $V' = 0$. Let

$$U_n' = \frac{1}{A_n} \sum_{k=1}^{n} a_k U_{n-k+1}, \qquad V_n' = \frac{1}{B_n} \sum_{k=1}^{n} b_k V_{n-k+1},$$

$$W_n' = \frac{1}{C_n} \sum_{k=1}^{n} c_k W_{n-k+1};$$

let $\sum_{k=1}^{\infty} u_k'$, $\sum_{k=1}^{\infty} v_k'$, and $\sum_{k=1}^{\infty} w_k'$ be the corresponding series; then

$$|W_n'| \leqq \frac{1}{C_n} \left\{ |u_1'| \left| \sum_{l=1}^{n} c_l V_{n-l+1} \right| + \sum_{k=2}^{n} \left[|u_k'| \left| \sum_{l=1}^{n-k} V_l S + A_k b_1 V_{n-k+1} \right| \right] \right\},$$

where

$$S = A_k b_{n-k-l+2} + \sum_{p=k+1}^{n-l+1} a_p b_{n-l-p+2}.$$

Hence

$$\begin{aligned} |W_n'| &\leqq \frac{1}{C_n} \left\{ |u_1'| \left| \sum_{l=1}^{n} c_l V_{n-l+1} \right| + \sum_{k=2}^{n} \left[|u_k'| A_k \left| \sum_{l=1}^{n-k+1} b_{n-k-l+2} V_l \right| \right] \right. \\ &\quad \left. + \sum_{k=2}^{n-1} \left[|u_k'| \left| \sum_{l=1}^{n-k} \sum_{p=k+1}^{n-l+1} a_p b_{n-l-p+2} V_l \right| \right] \right\} \\ &= P_n |u_1'| + \sum_{k=2}^{n} [|u_k'| Q_{nk}] + \sum_{k=2}^{n-1} [|u_k'| R_{nk}]. \end{aligned}$$

Since C includes B,*

$$\lim_{n\to\infty} P_n = 0. \tag{1}$$

For $2 \leqq k \leqq n$,

$$Q_{nk} < \frac{A_k \left| \sum_{l=1}^{n-k+1} b_{n-k-l+2} V_l \right|}{\sum_{l=k}^{n} A_l b_{n-l+1}}$$

$$< \frac{A_k \left| \sum_{l=1}^{n-k+1} b_{n-k-l+2} V_l \right|}{A_k \sum_{l=k}^{n} b_{n-l+1}} = \left| V'_{n-k+1} \right|.$$

Therefore

$$\sum_{k=2}^{n} Q_{nk} \left| u'_k \right| < \sum_{k=2}^{n} \left| V'_{n-k+1} \right| \left| u'_k \right|$$

$$= \sum_{k=2}^{\nu} \left| V'_{n-k+1} \right| \left| u'_k \right| + \sum_{k=\nu+1}^{n} \left| V'_{n-k+1} \right| \left| u'_k \right|,$$

where ν may be chosen so that ν and $n-\nu$ become infinite with n. Since $\lim_{n\to\infty} \left| V'_n \right| = 0$, for any ϵ, ν and n may be chosen sufficiently large so that

$$\sum_{k=2}^{n} [Q_{nk} \left| u'_k \right|] < \epsilon. \tag{2}$$

For $2 \leqq k \leqq n-1$, we have

$$R_{nk} \leqq \frac{\sum_{l=1}^{n-2} a_{n-l+1} B_l \left| V'_l \right|}{C_n}$$

$$= \left[\frac{1}{C_n} \sum_{l=1}^{n-2} a_{n-l+1} B_l \right] \left[\sum_{p=1}^{n-2} \frac{a_{n-p+1} B_p \left| V'_p \right|}{\sum_{l=1}^{n-2} a_{n-l+1} B_l} \right]$$

$$< \sum_{p=1}^{n-2} \frac{a_{n-p+1} B_p \left| V'_p \right|}{\sum_{l=1}^{n-2} a_{n-l+1} B_l} = V''_{n-2},$$

* Nörlund, Lunds Universitet, Årsskrift, (2), vol. 16 (1919), No. 3.

where V'_{n-2} is the $(n-2)$th term of the sequence obtained when $\{|V'_p|\}$ is summed by the matrix transformation t_{np}, where

$$t_{np} = \frac{a_{n-p+3}B_p}{\sum_{l=1}^{n} a_{n-l+3}B_l}$$

for $p \leqq n$, and $t_{np}=0$ for $p>n$.

This transformation is regular, since A is regular; it follows that $\lim_{n\to\infty} V''_{n-2}=0$, and that for any ϵ, we may find n sufficiently large so that

$$R_{nk} < \epsilon. \tag{3}$$

From (1), (2), (3), and the fact that $\sum_{k=1}^{\infty}|u'_k|$ converges, it follows that $\lim_{n\to\infty}|W'_n|=0$. This proves the theorem for $V'=0$.

If $V'\neq 0$, we consider the sequence $\{V_n - V'\}$; this sequence is summed by B to 0. Hence the Cauchy product of $\sum_{n=1}^{\infty}u_n$ by $[v_1 - V'] + \sum_{n=2}^{\infty}v_n$ is summed by C to 0; that is,

$$\lim_{n\to\infty}\left[W'_n - V'\frac{\sum_{k=1}^{n} c_k U_{n-k+1}}{C_n}\right] = 0.$$

Since C includes A,

$$\lim_{n\to\infty}\frac{\sum_{k=1}^{n} c_k U_{n-k+1}}{C_n} = U';$$

therefore $\lim_{n\to\infty} W'_n = U'V'$.

THEOREM 2. *If $\sum_{k=1}^{\infty}u_k$ is summable A to U', and if $\sum_{k=1}^{\infty}v_k$ is summable B to V', then $\sum_{k=1}^{\infty}w_k$ is summable D to $U'V'$.*

PROOF. We have

$$W'_n = \frac{1}{D_n}\sum_{k=1}^{n} d_k W_{n-k+1} = \sum_{k=1}^{n} g_{nk}U'_k V'_{n-k+1},$$

where $g_{nk}=A_k B_{n-k+1}/D_n$ for $k \leqq n$, and $g_{nk}=0$ for $k>n$.

Since A and B are regular, this method of summation is regular and $\lim_{n\to\infty} g_{n,n-k+1}=0$; it follows that*

$$\lim_{n\to\infty} \sum_{k=1}^{n} g_{nk} U_k' V_{n-k+1}' = U'V',$$

which completes the proof.

For the proof of Theorem 3 we require the following lemma.

LEMMA. *If* $\{X_n\}$ *and* $\{[\sum_{k=1}^{n} B_k y_{n-k+1}]/B_n\}$ *converge to* X *and* Y, *respectively, and if* $\{[\sum_{k=1}^{n} b_k y_{n-k+1}]/b_n\}$ *is bounded, then*

$$\lim_{n\to\infty} \left\{ \sum_{k=1}^{n} a_{nk} X_k \frac{\sum_{l=1}^{n-k+1} b_l y_{n-k-l+2}}{b_{n-k+1}} \right\} = XY,$$

provided that (a) $\lim_{n\to\infty} a_{nk}=0$; (b) $\sum_{k=1}^{n} |a_{nk}| < M$ *for all* n, *where* M *is a positive constant*; (c) T' *includes* T, *where* T' *and* T *are triangular matrix transformations defined by* $t'_{nk}=a_{n,n-k+1}$ *and* $t_{nk}=b_k/B_n$.

PROOF. Let

$$\frac{\sum_{l=1}^{n-k+1} b_l y_{n-k-l+2}}{b_{n-k+1}} = y'_{n-k+1},$$

and let $a_{nk}y'_{n-k+1}=c_{nk}$. Let $Z_n=\sum_{k=1}^{n} c_{nk}X_k$. From (a), (b), and (c), it follows that $\lim_{n\to\infty} c_{nk}=0$, $\sum_{k=1}^{n} |c_{nk}| < M'$, where M' is a positive constant, and $\lim_{n\to\infty} \sum_{k=1}^{n} c_{nk} = Y$. Choose p such that for a given $\epsilon>0$, $|X_k-X| < \epsilon/(2M')$ when $k>p$. For $k\leqq p$, $|X_k-X| < L$. Then

$$\left| Z_n - X\sum_{k=1}^{n} c_{nk} \right| \leqq \sum_{k=1}^{p} |c_{nk}||X_k - X| + \sum_{k=p+1}^{n} |c_{nk}||X_k - X|$$

$$\leqq L\sum_{k=1}^{p} |c_{nk}| + \frac{\epsilon}{2M'} |c_{nk}|.$$

Choose $N>p$, and such that $|c_{nk}| < \epsilon/(2pL)$ for $n>N$. Then for $n>N$, $|Z_n - X\sum_{k=1}^{n} c_{nk}| < \epsilon$, which proves the lemma.

---

* Dale, American Journal of Mathematics, vol. 47 (1925), p. 82.

THEOREM 3. *If* $\sum_{k=1}^{\infty} u_k$ *is summable A to U′ and* $\sum_{k=1}^{\infty} v_k$ *summable B to V′, and if*

$$\left| \frac{b_n v_1 + \cdots + b_1 v_n}{b_n} \right| < M,$$

then $\sum_{k=1}^{\infty} w_k$ *is summable C to U′V′.*

PROOF. Consider the triangular matrix definition

$$a_{nk} = \frac{A_k b_{n-k+1}}{A_1 b_n + \cdots + A_n b_1}.$$

This definition satisfies the three conditions of the lemma, for

$$(4) \qquad \frac{A_k b_{n-k+1}}{\sum_{l=1}^{n} A_l b_{n-l+1}} < \frac{A_k b_{n-k+1}}{\sum_{l=k}^{n} A_l b_{n-l+1}} < \frac{b_{n-k+1}}{\sum_{l=k}^{n} b_{n-l+1}};$$

$$(5) \qquad \sum_{k=1}^{n} |a_{nk}| = 1;$$

$$(6) \qquad A' = C'B',$$

where A', B', and C' are triangular matrix definitions with

$$a'_{nk} = a_{n,n-k+1}, \qquad b'_{nk} = \frac{b_k}{B_n}, \qquad c'_{nk} = \frac{a_{n-k+1} B_k}{\sum_{l=1}^{n} A_l b_{n-l+1}}.$$

The definition C' is regular. The theorem follows immediately from this lemma.

THE GEORGE WASHINGTON UNIVERSITY

ON THE LOCUS OF AN ANALYTIC EQUATION IN THE REAL PLANE*

BY A. B. BROWN

I have been unable to find in the literature a statement or proof of the following theorem.

THEOREM. *Let $f(x, y)$ be analytic at the real point (x_0, y_0), with $f(x_0, y_0)=0$ and $f(x, y)$ irreducible at (x_0, y_0).† Then the locus of the equation $f(x, y)=0$ in the real xy plane near (x_0, y_0) consists of one of the following three*: (1) *the point (x_0, y_0)*; (2) *a single smooth curve through (x_0, y_0)*; (3) *a cusp with vertex at (x_0, y_0).*

More detailed descriptions of (2) and (3) are contained in the proof which follows.

By a change of coordinates we may suppose $x_0=y_0=0$. According to the Weierstrass preparation theorem for the case of one independent variable, since f is irreducible at (0, 0), either $f(x, y)\equiv x\Omega(x, y)$, with $\Omega(0, 0)\neq 0$ and Ω analytic at (0, 0) or

$$f(x, y) \equiv [y^m + A_1(x)y^{m-1} + \cdots + A_m(x)]\Omega(x, y), \tag{1}$$

with Ω as above, $m>0$, $A_j(x)$ analytic at $x=0$, and $A_j(0)=0$, $(j=1, \cdots, m)$. Since in the first case the real locus $f(x, y)=0$ is merely a straight line, it is sufficient to consider the case that (1) holds.

Since f is irreducible at (0, 0), the same is true of the algebroid function in (1), and hence its m-leaved Riemann surface is connected near (0, 0) and we can uniformize locally as follows:

$$x = t^m, \tag{2}$$

$$y = \psi(t) = a_1t + a_2t^2 + \cdots, \tag{3}$$

with ψ analytic at $t=0$, and a neighborhood of the origin in the

* Presented to the Society, February 23, 1935.

† That is, not the product, near (x_0, y_0) in the 4-space of the complex variables, of two functions each analytic and zero at (x_0, y_0). For theorems which we use involving functions of complex variables, see W. F. Osgood, *Lehrbuch der Funktionentheorie*, vol. 1, Chapter 8, §14, and vol. 2, part 1, Chapter 2, §§2, 4, 7, 9, 10, 11.

t plane giving exactly the points (x, y) in a neighborhood of $(0, 0)$ satisfying $f(x, y) = 0$, each just once. Now x is real if and only if

$$t = \tau e^{k\pi i/m}, \qquad (k = 0, 1, \cdots, m-1), \tag{4}$$

with τ real and positive, negative, or zero. Substituting from (4) in (3), for a fixed value of k we get

$$y = \phi_k(\tau) = b_{k,1}\tau + b_{k,2}\tau^2 + \cdots, \qquad b_{k,\nu} = a_\nu e^{k\nu\pi i/m}. \tag{5}$$

Either all the coefficients in (5) are real, in which case *equation* (4) *for the given* k, *together with* (2) *and* (3), *gives only real points* (x, y), or else there is a first non-real $b_{k,j}$. In this case we have from (5) that

$$y = R(\tau) + \tau^j(b_{k,j} + b_{k,j+1}\tau + b_{k,j+2}\tau^2 + \cdots),$$

where $R(\tau)$ is real. Since $b_{k,j}$ is not real, we see that for sufficiently small τ the parenthesis is not real. Since τ is real, it follows that for small values of $\tau \neq 0$, y is not real. When $\tau = 0$, $y = 0$. We thus see that in this case *only* $\tau = 0$ *gives real* y. Next we show that at most one value of k can give real y for real $\tau \neq 0$.

Suppose $t = \tau e^{k\pi i/m}$ and $t = \tau e^{j\pi i/m}$, with $0 < j - k < m$, both gave real loci for real τ, when used with (2) and (3). Then $b_{k,\nu} = a_\nu e^{k\nu\pi i/m}$ and $b_{j,\nu} = a_\nu e^{j\nu\pi i/m}$ would be real, $(\nu = 1, 2, \cdots)$, and hence if $a_\nu \neq 0$, their ratio $e^{(j-k)\nu\pi i/m}$ would be real. Therefore $(j-k)\nu/m$ would be an integer. Let $m = m_1 \cdot m_2$, where m_1 is the H. C. F. of $(j-k)$ and m. Thus $m_2 > 1$, since $0 < j - k < m$. Then ν would have to be a multiple of m_2, and hence

$$y = \psi(t) \equiv \zeta(t^{m_2}), \tag{6}$$

with ζ analytic at the origin. Now if we set $t_1 = te^{2\pi i/m_2}$, then $(t_1)^m = t^m e^{2\pi m_1 i} = t^m$, and $t_1{}^{m_2} = t^{m_2} e^{2\pi i} = t^{m_2}$. Hence from (2) and (6) we see that t_1 and t would give the same (x, y). But since $m_2 > 1$, $t_1 \neq t$ if $t \neq 0$, and therefore we would have a contradiction. Consequently *at most one of the* m *values of* k *gives real* (x, y) *when used in* (2), (3), *and* (4) *with real* $\tau \neq 0$.

We now consider the case that some value of k does give a real locus. We define as the *branches* of the locus for the k in question the parts for which $\tau \geqq 0$ and $\tau \leqq 0$, respectively. Thus the two branches have the origin as a common end point. From (4), (2), and (3) we see that *on each of the two branches* y *is a*

single-valued real function of x, analytic when $x \neq 0$. If m is even, x has the same sign on each branch. If m is odd, x has opposite signs on the two branches except at $x=0$, and y is a single-valued real continuous function of x for x on an interval of the real axis containing the origin.

From (2) and (3) we see that (i) if $a_1=a_2=\cdots=a_m=0$, then $\lim_{x\to 0} y/x=0$; (ii) if $a_1=\cdots=a_{m-1}=0$ but $a_m\neq 0$, then $\lim_{x\to 0} y/x=a_m\neq 0$; (iii) if some $a_j\neq 0$ with $0<j<m$, then when x approaches zero, y/x becomes infinite. Hence the two branches have a common tangent line. By a rotation of axes we now arrange that the x axis is the tangent line at the origin, and that $x\geqq 0$ on the real locus if m is even. Since $\lim_{x\to 0} y/x$ must then be 0, we must have case (i), so that $a_1=\cdots=a_m=0$. Then, for $t\neq 0$,

$$(7)\qquad \frac{dy}{dx}=\left(\frac{dy}{dt}\right)\left(\frac{dt}{dx}\right)=[(m+1)a_{m+1}t^m+\cdots]\left(\frac{1}{mt^{m-1}}\right).$$

Therefore $\lim_{x\to 0} dy/dx=0$. Hence we have a *continuously turning tangent*; though if m is even, properly speaking we may only say that each branch has the positive x axis as right-hand tangent at the origin (a cusp).

Now from (7), for $t\neq 0$,

$$(8)\qquad \frac{dy}{dx}=b_{m+1}t+b_{m+2}t^2+\cdots.$$

Hence when $t\neq 0$,

$$(9)\qquad \frac{d^2y}{dx^2}=\left[\frac{d}{dt}\left(\frac{dy}{dx}\right)\right]\cdot\frac{dt}{dx}=(b_{m+1}+2b_{m+2}t+\cdots)\left(\frac{1}{mt^{m-1}}\right).$$

Therefore on either branch, with $t\neq 0$, d^2y/dx^2 is either identically zero or nowhere zero. Hence *unless the locus is a straight line, d^2y/dx^2 has a fixed sign on each branch except at (x_0, y_0) where it may be zero, and the branch does not meet T, the tangent at (x_0, y_0), except at (x_0, y_0).* We also see that *as a point on the locus approaches (x_0, y_0), the curvature either approaches a fixed limit, possibly zero, or becomes infinite. The curvature at (x_0, y_0) exists for each branch and equals this limit* (also in the infinite case, where we have infinite curvature at (x_0, y_0)), as follows from (2), (8), and (9).

As examples we mention $y^3-x^4=0$, $y^3-x^5=0$, and $[y-1+(1-x^2)^{1/2}]^2-x^5=0$ with the determination of $(1-x^2)^{1/2}$ which equals 1 when $x=0$. These give respectively, at the origin, a minimum, a point of inflection, and a cusp with both branches concave upward. In none of the three cases is y analytic in x at the origin. An example where the locus is a single point is given by $y+ix=0$.

In the case of a reducible function $f(x, y)$, the real locus $f(x, y)=0$ neighboring (x_0, y_0) consists of a finite number of configurations of the kind described in the theorem, no two of which have any point except (x_0, y_0) in common. This is easily proved by use of theorems on resultants and on divisibility of one function by another. Of course two irreducible factors may have exactly the same locus.

COLUMBIA UNIVERSITY

A PARTIAL DIFFERENTIAL EQUATION CONNECTED WITH THE FUNCTIONS OF THE PARABOLIC CYLINDER*

BY HARRY BATEMAN

The partial differential equation

$$\sum_{s=1}^{p}\left(\frac{\partial^2 V}{\partial x_s^2}-x_s\frac{\partial V}{\partial x_s}\right)+\nu V=0, \tag{1}$$

which was considered by Mehler† in 1866, is a slight modification of an equation which occurs in wave-mechanics in the theory of the rotator in a plane and in space.‡ The case in which ν is a positive integer is then of chief physical interest and Mehler's simple solution

$$V=\prod_{s=1}^{p}H_{m_s}(x_s), \qquad \sum_{s=1}^{p}m_s=\nu, \tag{2}$$

acquires a physical significance. The function $H_m(x)$ is the polynomial of Laplace and Hermite defined by the equation

* Presented to the Society, December 2, 1933.

† F. G. Mehler, Journal für Mathematik, vol. 66 (1866), p. 161.

‡ A. Sommerfeld, *Atombau und Spektrallinien, wellenmechanischer Ergänzungsband*, 1929, p. 23.

$$H_m(x) = e^z(-D_z)^m e^{-z} = (2\pi)^{1/2} e^z \phi^{(m)}(x), \tag{3}$$

where $2z = x^2$. The H-notation is that adopted by Appell and de Fériet in their book,* the ϕ-notation is one used in statistics.

When $p = 2$ we have the equation

$$\frac{\partial^2 V}{\partial x^2} + \frac{\partial^2 V}{\partial y^2} - x\frac{\partial V}{\partial x} - y\frac{\partial V}{\partial y} + \nu V = 0, \tag{4}$$

which is closely related to one considered by E. L. Ince.† The *normal* or simple solutions are of type

$$V = h_{r-m}(x) k_m(y), \tag{5}$$

where $h_s(x)$, $k_s(x)$ are solutions of the differential equation

$$\frac{d^2u}{dx^2} - x\frac{du}{dx} + su = 0. \tag{6}$$

Ince generalized this solution by multiplying it by an arbitrary function $f(n)$ and integrating with respect to n. We shall use summation instead of integration and shall suppose $m+1$ to be a positive integer. We shall, moreover, put $h_s(x) = k_s(x) = H_s(x)$, where $H_s(x)$ is defined for one-half of the x plane and for all values of s by the equation

$$C_{m+s,m} m! i^m H_s(x) = \int_0^\infty [(x + iy)^{s+m}\phi^{(m)}(y) + (x - iy)^{s+m}\phi^{(m)}(-y)]dy, \tag{7}$$

which reduces to the well known formula of Laplace‡ and Mehler§ when $m = 0$ and s is a positive integer. On the right-hand side of the equation (7), the many-valued function $(x \pm iy)^{2n}$ is supposed to be defined for $y > 0$ so that when $x \to 0$ it reduces to $y^{2n}\exp(\pm in\pi)$. When x is complex and unrestricted the integral is not single-valued, but if x is restricted to the half-plane

* P. Appell and J. Kampé de Fériet, *Fonctions Hypergéométriques et Hypersphériques, Polynomes d'Hermite*, 1926, p. 344.

† E. L. Ince, Proceedings of the Royal Society of Edinburgh, vol. 44 (1924), p. 242.

‡ P. S. Laplace, *Théorie Analytiques des Probabilités*, Livre 2, Oeuvres, vol. 7, p. 299.

§ F. G. Mehler, loc. cit.

$R(x)>0$, the integral agrees in value with a single-valued function which satisfies the differential equation (6). It is, indeed, easy to verify by an integration by parts that the definite integral has the form on the left, where $H_s(x)$ is some function of s and x such that $H_s(x)$ satisfies (6). The proof of the last property depends on the fact that $(x+iy)^r$ is a solution of the differential equation (4).

To find which solution of (6) the integral represents we first choose $R(s+m)>0$ and make $x\to 0$. Writing $s=2r$ and observing that*

$$(8)\quad 2\pi\int_0^\infty y^{m+2r}\phi^{(m)}(y)\,dy = C_{2r+m,m}\,m!\,2^r\Gamma\left(\frac{1}{2}\right)\Gamma\left(r+\frac{1}{2}\right),$$

we find that

$$(9)\quad \begin{aligned} H_{2r}(0)\Gamma\left(\frac{1}{2}-r\right) &= 2^r\Gamma\left(\frac{1}{2}\right), \\ H_{2r}'(0)\Gamma(-r) &= 2^{r-1/2}\Gamma\left(-\frac{1}{2}\right). \end{aligned}$$

Hence, with the usual notation for the confluent hypergeometric function (without the suffixes),

$$(10)\quad \begin{aligned} H_{2r}(x)\Gamma(-r) = {} & 2^r B\left(\frac{1}{2},-r\right)F\left(-r;\frac{1}{2};\frac{x^2}{2}\right) \\ & + \Gamma\left(-\frac{1}{2}\right)2^{r-1/2}\cdot xF\left(\frac{1}{2}-r;\frac{3}{2};\frac{x^2}{2}\right). \end{aligned}$$

In Whittaker's notation, if s is not a negative integer,

$$(11)\quad H_s(x) = \exp\left(\frac{x^2}{4}\right)D_s(x).$$

This equation may be used to define the function $H_s(x)$ for all real and complex values of x. An equation equivalent to this has

* The value of the integral is given in a slightly different form by G. N. Watson, Proceedings of the London Mathematical Society, vol. 8 (1910), p. 393.

indeed been used for this purpose* for all real and complex values of s except negative integral values. When $R(s+m)>0$ we can, by integrating the integral in (7) a sufficient number of times, pass to the case in which $R(s+m)<0$. An exceptional case arises when s is a negative integer. If we take $m+s$ to be a negative integer, the integral has a definite value and gives a definition of $H_s(x)$, but a different notation is used for the function thus defined.† If $m+s$ is a positive integer, the integral and $C_{m+s,\,m}$ are both zero for $s<0$. When s is a positive integer $p-n$, the relation (7) gives at once the expansion

$$(12) \qquad (x+iy)^p = \sum_{m=0}^{p} C_{p,m} i^m H_{p-m}(x) H_m(y),$$

which may be checked by comparing the terms of degree p on the two sides of the equation and noting that both sides satisfy the partial differential equation (4). This equation may be used to derive from (7) the expansion (s not a negative integer)

$$(13) \qquad H_{s+p}(x) = \sum_{m=0}^{p} (-1)^m m! C_{p,m} C_{s,m} H_{s-m}(x) H_{p-m}(x),$$

which may be inverted so as to give an expansion for the product in the form‡

$$(14) \qquad H_s(x) H_n(x) = \sum_{m=0}^{n} m! C_{s,m} C_{n,m} H_{n+s-2m}(x).$$

This relation may be established directly by induction.§ When $s=n$ the equation takes the form

$$(15) \qquad [H_n(x)]^2 = \sum_{s=0}^{n} C_{n,s}\, n!\, (s!)^{-1} H_{2s}(x).$$

---

* M. Plancherel and W. Rotach, Commentarii Mathematici Helvetici, vol. 1 (1929), p. 227. See also S. C. van Veen, Amsterdam Proceedings, vol. 34 (1931), p. 257.

† G. N. Watson, loc. cit.; O. Volk, Mathematische Annalen, vol. 86 (1922), p. 296; Appell and de Fériet, loc. cit., p. 362.

‡ An equivalent expansion is given by Gorakh Prasad for positive integral values of s and n. Products of Hermite polynomials are considered also by N. Nielsen, *Recherches sur les Polynomes de Hermite*, Copenhagen, 1918.

§ Gorakh Prasad, Proceedings of the Benares Society, vol. 2 (1920), p. 18.

This series may be compared with Mitra's series* for $D_n^2(x)$,

$$(16)\quad [D_n(x)]^2=\sum_{s=0}^{\infty}(2s-1)!!(2n-2s-1)!![(2s)!!2^{n+s+1/2}]^{-1}H_{2s}(x),$$

which has been written in an abbreviated form with the aid of Schuster's notation†

$$(2s)!! = (2s)(2s-2)\cdots(2),$$
$$(2s+1)!! = (2s+1)(2s-1)\cdots(1),$$

and a natural extension of it, namely,

$$(-1-2s)!!(-1)(-3)\cdots(1-2s) = 1.$$

Mitra's expansion may be generalized with the aid of the well known recurrence relations

$$(17)\quad D_{n+1}(x) - xD_n(x) + nD_{n-1}(x) = 0,$$

$$(18)\quad D_n'(x) + \frac{1}{2}xD_n(x) - nD_{n-1}(x) = 0,$$

$$(19)\quad H_{n+1}(x) - xH_n(x) + nH_{n-1}(x) = 0,$$

$$(20)\quad H_n'(x) = nH_{n-1}(x).$$

The resulting equations are

$$(21)\quad D_n(x)D_{n-2r}(x) = \sum_{s=0}^{\infty}(2r+2s-1)!!(2n-2r-2s-1)!!(2s-2r-1)!!A_s,$$

where

$$A_s\cdot(2s)!!(2s-1)!!2^{n-r+s+1/2} = H_{2s}(x);$$

$$(22)\quad D_n(x)D_{n-2r+1}(x) = \sum_{s=0}^{\infty}(2r+2s-1)!!(2n-2r-2s-1)!!(2s-2r+1)!!C_s,$$

where

* S. C. Mitra, Proceedings of the Edinburgh Society, (2), vol. 4 (1934), p. 27.

† A. Schuster, Philosophical Transactions, Royal Society of London, (A), vol. 200 (1903), p. 181. The extension is also given by Schuster.

$$C_s(2s)!!(2s+1)!!2^{n-r+s+3/2} = H_{2s+1}(x).$$

These equations may be checked* by comparing them with (13) and (14) after these have been multiplied throughout by $D_0^2(x)$ so that they become equations between D-functions. The verification depends on the identities

$$\begin{aligned}
&(2m+2n-1)!!(-2n-1)!!(2r-1)!! \\
&= \sum_{s=0}^{m} 2^s C_{m,s} C_{m+n+r,s}\, s!\,(2m+2r-2s-1)!! \\
&\qquad \cdot (2s-2m-2n-1)!!(2m+2n-2s-1)!!; \\
&(2n+2s-1)!!(2n-2s-1)!!(2s-2n-1)!! \\
&= \sum_{m=0}^{p} (-2)^m C_{p,m} C_{2n-p,m}\, m!\,(2p-2n+2s-1)!! \\
&\qquad \cdot (2n-2m-2s-1)!!(2s+2n-2p-1)!!.
\end{aligned} \tag{23}$$

Let us now consider the integral

$$H_{\nu,m}(\alpha, x) = \int_{-\infty}^{\infty} H_\nu(x\cos\alpha + y\sin\alpha)\phi^{(m)}(y)dy, \tag{24}$$

which is convergent when $-\pi<2\alpha<\pi$ if $H_\nu(z)$ is defined by (11), for the behavior of $D_\nu(z)$ when z is large is known. Since

$$\frac{\partial}{\partial z} H_\nu(z) = \nu H_{\nu-1}(z),$$

we readily find that

$$\begin{aligned}
\frac{\partial}{\partial\alpha} H_{\nu,0}(\alpha, x) &= -x\sin\alpha \int_{-\infty}^{\infty} \nu H_{\nu-1}(\ )\phi^{(0)}(y)dy \\
&\quad + \cos\alpha \int_{-\infty}^{\infty} \nu H_{\nu-1}(\ ) y\phi^{(0)}(y)dy \\
&= -\nu x\sin\alpha H_{\nu-1,0}(\alpha, x) \\
&\quad + \nu(\nu-1)\cos\alpha\sin\alpha H_{\nu-2,0}(\alpha, x).
\end{aligned} \tag{25}$$

* The expansions obtained by putting $n=2r$ in (21) and $n=2r-1$ in (22) have already been given by R. K. Varma in a slightly different form, Proceedings of the Benares Society, vol. 10 (1928), p. 11. Some special cases were considered previously by Gorakh Prasad, Proceedings of the Benares Society, vol. 2 (1920), p. 12.

Since, when $\alpha \to 0$,

$$H_{\nu,0}(0, x) = H_\nu(x) = xH_{\nu-1}(x) - (\nu - 1)H_{\nu-2}(x), \tag{26}$$

we find that a solution of equation (25) is

$$H_{\nu,0}(\alpha, x) = \cos^\nu \alpha H_\nu(x). \tag{27}$$

Integrating the integral (24) by parts m times we find

$$H_{\nu,m}(\alpha, x) = C_{\nu,m} m! \cos^{\nu-m} \alpha \sin^m \alpha H_{\nu-m}(x). \tag{28}$$

The associated equation*

$$H_\nu(x \cos \alpha + y \sin \alpha) = \sum_{m=0}^{\infty} C_{\nu,m} \cos^{\nu-m} \alpha \sin^m \alpha H_{\nu-m}(x) H_m(y) \tag{29}$$

was obtained by Kampé de Fériet† in the case when ν is a positive integer by using the generating function of the Hermite polynomial, a method which he used also to obtain an expansion for $H(p_1x_1+p_2x_2+ \cdots +p_sx_s)$ in a series of simple functions of type (1) on the supposition that $p_1^2+p_2^2+ \cdots =1$. In the general case when ν is not a positive integer we may consider the series in powers of $\tan \alpha$ of $\sec^\nu \alpha H_\nu(x \cos \alpha + y \sin \alpha)$. If ν is not a negative integer, the radius of convergence of this series is determined by the location of the singularities of $\sec^\nu \alpha$, for $H_\nu(z)$ is uniform and analytic over the whole of the finite part of the z-plane. The power series for $\sec^\nu \alpha H_\nu(\ )$ is thus convergent for $-\pi < 2\alpha < \pi$, and so to establish the truth of the series under this condition it is only necessary to prove that if $t = \tan \alpha$, then

$$I(m, \nu) = \left[\frac{d^m}{dt^m} \sec^\nu \alpha H_\nu(x \cos \alpha + y \sin \alpha)\right]_{t=0} = C_{\nu,m} m! H_{\nu-m}(x) H_m(y). \tag{30}$$

* When ν is a positive integer this equation may be checked by observing that both sides are polynomial solutions of (4) with the same terms of degree ν and these terms determine the solution uniquely.

† J. Kampé de Fériet, Danske Videnskabernes Selskabs, Mathematisk-Fysiske Meddelelser, vol. 5 (1923), No. 2. See also Appell and de Fériet, loc. cit., and V L. Mutatker, Journal of the Indian Mathematical Society, vol. 1 (1934), p. 53. The case in which $\alpha = \pi/4$ is due to C. Runge, Mathematische Annalen, vol. 75 (1914), p. 130. The equation was in fact, used by him to solve a special type of non-linear integral equation.

This proof may be carried out by induction. Thus, if we suppose that $\sec^\nu \alpha H_\nu(x\cos\alpha+y\sin\alpha)=F(\nu,t)$, we have, on differentiating first once and then m times,

$$(1+t^2)\frac{d}{dt}F(\nu,t)=(y-xt)\nu F(\nu-1,t)+\nu t F(\nu,t),$$

$$I(m+1,\nu)=y\nu I(m,\nu-1)-mx\nu I(m-1,\nu-1)+m(\nu-m-1)I(m-1,\nu),$$

for $t=0$, and the recurrence relation is satisfied by the expression on the extreme right of (30).

When ν is a positive integer, the integral representing $H_{\nu,0}(\alpha,x)$ may be evaluated by putting $z=x\cos\alpha+y\sin\alpha$ and using Mehler's expansion*

$$\csc\alpha\cdot\exp\left[\csc^2\alpha(x^2\cos^2\alpha-2xz\cos\alpha+z^2\cos^2\alpha)\right]$$

$$=\sum_{s=0}^{\infty}\frac{1}{s!}\cos^s\alpha\cdot H_s(x)H_s(z). \tag{31}$$

When the integral for $H_{\nu,m}(\alpha,x)$ is treated in the same way, we obtain the integral

$$\int_{-\infty}^{\infty}H_\nu(z)\phi_m(z\csc\alpha-x\operatorname{ctn}\alpha)dz$$

$$=C_{\nu,m}m!\cos^{\nu-m}\alpha\sin^m\alpha H_{\nu-m}(x), \tag{32}$$

which gives the expansion

$$e^{z^2/2}\phi_m(z\csc\alpha-x\operatorname{ctn}\alpha)$$

$$=(2\pi)^{-1/2}\sum_{n=m}^{\infty}[(n-m)!]^{-1}\cos^{n-m}\alpha\sin^m\alpha H_n(z)H_{n-m}(x). \tag{33}$$

Another expansion may be obtained by writing $\operatorname{ctn}\alpha=e^{-t}$ and observing that the function $V=\csc^\nu\alpha\cdot H_\nu(x\cos\alpha+y\sin\alpha)$ is a solution of the partial differential equation†

---

* F. G. Mehler, loc. cit.

† This equation may be transformed into one which Laplace solved by means of a definite integral, *Théorie Analytique des Probabilités*, Livre 2, Oeuvres, vol. 7, p. 294. The equation was used later by Smoluchowski in his theory of the Brownian movement. See B. Hostinský, Annales de l'Institut Poincaré, vol. 3 (1932), pp. 1–72.

$$\frac{\partial^2 V}{\partial x^2} - x\frac{\partial V}{\partial x} = \frac{\partial V}{\partial t}. \tag{34}$$

If this equation has a solution of the form

$$V = \sum_{k=0}^{\infty} C_k(t)x^k/k!, \tag{35}$$

the coefficient $C_k(t)$ must satisfy the differential-difference equation

$$C_k'(t) + kC_k(t) = C_{k+2}(t), \tag{36}$$

and if the value of $C_k(0)$ is known, an appropriate solution is

$$C_k(t) = \sum_{s=0}^{\infty} 2^{-s}C_{k+2s}(0)e^{-kt}(1 - e^{-2t})^s/s!. \tag{37}$$

Instead of defining $C_k(t)$ as a coefficient we may define it by the equation

$$C_k(t) = \left(\frac{\partial^k V}{\partial x^k}\right)_{x=0}. \tag{38}$$

In particular, when $V = \csc^\nu \alpha \cdot H_\nu(x\cos\alpha + y\sin\alpha)$, we have

$$C_k(t) = C_{\nu,k}\, k! \cos^k \alpha \csc^\nu \alpha H_{\nu-k}(y\sin\alpha), \tag{39}$$

and if $n = \nu - k$, the foregoing expansion for $C_k(t)$ leads to the equation

$$H_n(y\sin\alpha) = \sum_{s=0}^{\infty} C_{n,s}H_{n-s}(y(2)^{1/2})H_s(0)2^{(n-s)/2}\sin^{n-s}\alpha\cos^s\alpha, \tag{40}$$

which may be rewritten in the form

$$H_n(uv) = \sum_{s=0}^{\infty} C_{n,s}H_{n-s}(u)H_s(0)v^{n-s}(1 - v^2)^{s/2}. \tag{41}$$

When n is a positive integer, this expansion is equivalent to one given by Appell and de Fériet.* The associated integral

* Appell and de Fériet, loc. cit., p. 346. Equation (40) is related also to equation (67) in the paper of S. Goldstein, Proceedings of the London Mathematical Society, (2), vol. 34 (1932), p. 103.

$$(42) \qquad \int_{-\infty}^{\infty} H_m(uv)\phi_{m-s}(u)du = m!H_s(0)v^{m-s}(1-v^2)^{s/2}/s!$$

is closely connected with one given by Doetsch.*

If in the expansion (29) we put $\alpha=\pi/4$ and set first $x=u(2)^{1/2}$, $y=v(2)^{1/2}$, secondly $x=(u+v)(2)^{1/2}$, $y=0$, we obtain two expansions for $H_\nu(u+v)$. Equating them and writing x for $u(2)^{1/2}$, y for $v(2)^{1/2}$, we obtain the equation (for $\nu>0$)

$$\sum_{m=0}^{\infty} C_{\nu,m}H_{\nu-m}(x)H_m(y) = \sum_{m=0}^{\infty} C_{\nu,m}H_{\nu-m}(x+y)H_m(0).$$

This was obtained by Varma† for the case in which ν is a positive integer. When $x=y$ Runge's expansion

$$2^\nu H_\nu(x(2)^{1/2}) = \sum_{m=0}^{\infty} C_{\nu,m}H_{\nu-m}(x)H_m(x)$$

suggests that it may be profitable to study the expansion of a product $H_r(x)H_s(x)$ in a series of functions of type $H_n(x(2)^{1/2})$. Such expansions have been studied by Mitra.‡

CALIFORNIA INSTITUTE OF TECHNOLOGY

* G. Doetsch, Mathematische Zeitschrift, vol. 32 (1930), p. 587.

† R. S. Varma, Proceedings of the Benares Society, vol. 9 (1927), p. 31.

‡ S. C. Mitra, Bulletin of the Calcutta Mathematical Society, June, 1926; Proceedings of the Benares Mathematical Society, vol. 9 (1927), p. 21; Proceedings of the Edinburgh Mathematical Society, (2), vol. 4 (1934), p. 27. See also S. C. Dhar, Journal of the Indian Mathematical Society, new ser., vol. 1 (1934), p. 105.

A GENERALIZATION OF THE BERNOULLI POLYNOMIAL OF ORDER ONE*

BY B. F. KIMBALL

1. *Introduction.* The idea of a Bernoulli polynomial, generalized so as to allow of a continuous variation of the index of its degree, is not a new one.† In the present paper the writer has developed a simple definition of a generalized Bernoulli polynomial of order one which brings the Bernoulli polynomial into direct relationship with the generalized Riemann zeta-function. Another very interesting property of this generalization of the Bernoulli polynomial is brought out in §7.

2. *Difference Equation Considered.* For the purposes of this paper in dealing with a complex power s of a complex number t, if $t=re^{i\theta}$, then $t^s \equiv t^{\sigma+i\tau} = e^{\sigma[\log r+i\theta]}$, $(-\pi<\theta\leqq\pi)$. The following *difference equation* will frequently be referred to:

$$\frac{1}{w}\left[f(x+w, s) - f(x, s)\right] = sx^{s-1}, \tag{1}$$

where the difference interval w is taken real and positive. Also there will be occasion to impose the *asymptotic condition: for any value of s such that* $R(s)<0$,

$$f(x, s) \to 0 \quad as \quad R(x) \to +\infty. \tag{2}$$

There will also be occasion to refer to the following regions on the x and s planes:

Region X. All points on the x plane other than the negative axis of reals and the origin.

Region S. All points on the s plane within and on the boundary of a circle of radius M centering at the origin, (usually taken arbitrarily large, see §5).

3. *Uniqueness of an Analytic Solution.*

---

* Presented to the Society, June 23, 1933.

† See N. E. Nörlund, *Vorlesungen über Differenzenrechnung*, p. 53. Future references to this book will be indicated by the letter N.

LEMMA 1. *Given a function of x and s which is defined for all values of x on region X and which for any such value of x is analytic in s over region S. Given also that this function is a solution of the difference equation* (1) *for x in X and that it satisfies the asymptotic condition* (2). *Such a function is unique* (the existence of such a function is established in §5).

PROOF. Assuming the existence of two such functions $f_{1,2}(x, s)$, we have

$$\psi(x + w, s) - \psi(x, s) = 0, \qquad (x \subset X; \psi \equiv f_1 - f_2). \tag{3}$$

From (2) it follows that $\psi(x, s)$ must be identically zero, $x \subset X$, if $R(s) < 0$. Let x_0 denote any value of x in region X. We have then from the hypothesis of Lemma 1 that $\psi(x_0, s)$ is an analytic function of s over the region S. Thus this analytic function must be zero for all values of s in S, and the lemma follows.

4. *Definition of a Generalized Bernoulli Polynomial.* For s = a positive integer or zero and x real, the classical Bernoulli polynomial, $B_s(x) \equiv B_s(x, 1)$, has been defined (N., p. 18) as the polynomial solution of the difference equation (3), with difference interval equal to unity, which takes on the value of the corresponding Bernoulli number B_s when $x = 0$. Nörlund has termed the polynomial above the "Bernoulli polynomial" (of order one), with difference interval equal to one, and has defined Bernoulli polynomials of positive and negative order n for all integral values of n, where the difference interval w is not restricted to be equal to unity (N., pp. 129, 138).

The author has found it of interest to set up a definition of a generalized Bernoulli polynomial $B_s(x, w)$ of order one where s and x are allowed to take on complex values. To avoid complications, w has been taken as real and positive.

DEFINITION. *The solution of the difference equation* (1) *described in the statement of Lemma* 1 *shall be considered as the generalized Bernoulli polynomial $B_s(x, w)$ of order one for s in region S.*

5. *Solution Based on Euler-Maclaurin Summation Formula.* A solution based upon the Euler-Maclaurin summation formula will first be set up. We define a periodic function $\overline{B}_m(t)$ as follows (N., p. 30). Let the variable t be restricted to real values

and denote by $B_m(t)$ the classical Bernoulli polynomial of degree m. The function $\overline{B}_m(t)$ is taken so that

$$\overline{B}_m(t) \equiv B_m(t), \qquad (0 \leqq t < 1), \tag{4}$$

and $\overline{B}_m(t+1) - \overline{B}_m(t) = 0$ for all values of t. Let m be any positive integer greater than a positive number M (radius of circular region S). Define the function $f(x, s) \equiv F(x+hw, s)$ by the relation (see N., p. 52):

$$\begin{aligned} F(x + hw, s) &= \sum_{k=0}^{m} w^k B_k(h) C_{s,k} x^{s-k} \\ &+ (m+1) w^{m+1} C_{s,m+1} \int_0^\infty \overline{B}_m(h - z)(x + wz)^{s-m-1} dz, \\ &\qquad (0 \leqq h < 1). \end{aligned} \tag{5}$$

It is not difficult to verify the following statements concerning this function for x in region X and s in region S.

(i) *For any fixed value of x in region X, $f(x, s)$ exists and is an analytic function of s in S.*

(ii) *The function $f(x, s)$ defined above is a solution of the difference equation* (1).

This is easily verified as follows. Consider the infinite integral as a sum of an infinite series of integrals with limits at $0, w, 2w, \cdots$. Then take the difference of each side of the equation with respect to x with difference interval w. This reduces the infinite series of integrals to one integral, and the expression on the right is easily recognized to be the Euler-Maclaurin expansion of $s(x+hw)^{s-1}$ in terms of its first differences.

(iii) *For any value of s such that $R(s) < 0$, $f(x, s) \to 0$ as $R(x) \to +\infty$.*

(iv) *The function $f(x, s)$ is the generalized Bernoulli polynomial, $B_s(x, w)$, of order one for s in region S.*

(v) *If $w = 1$ and s is a positive integer (or zero) less than M, $f(x, s)$ reduces to the classical Bernoulli polynomial.*

(vi) The above region of definition S of $B_s(x, w)$ on the s plane may be made as large as desired by giving to m a sufficiently large value in equation (5). Thus *we have here a proof*

of the existence of $B_s(x, w)$ and its analyticity with respect to s over the finite s plane for $x \subset X$.

(vii) Set $h=0$ and $x=w$ in (5). This brings out the relation (N., p. 129)

$$B_s(w, w) = w^s B_s(1, 1) \equiv w^s B_s(1). \tag{6}$$

6. *Multiplication Equation Satisfied by* $B_s(x, w)$.* The difference equation

$$h_s(x + w/n) - h_s(x) = wsx^{s-1}, \qquad (x \subset X),$$

is satisfied by $\sum_{k=0}^{n-1} B_s(x+kw/n, w)$ and $n^{1-s}B_s(nx, w)$. When $R(s)<0$ the asymptotic value of each of these functions, as $R(x)$ becomes positively infinite, is zero. Thus, with the reasoning employed in the proof of Lemma 1, it follows that

$$n^{1-s}B_s(nx, w) = \sum_{k=0}^{n-1} B_s(x + kw/n, w), \qquad (x \subset X), \tag{7}$$

for all values† of s. In particular for $n=2$ and $x=w/2$, this gives the relation

$$B_s(w/2, w) = (2^{1-s} - 1)B_s(w, w),$$

which taken in connection with (6), leads to

$$B_s(w/2, w) = w^s B_s(1/2). \tag{8}$$

7. *An Interesting Property of $B_s(x, w)$.* Define function $\phi_s(x)$ by the relation:

$$\phi_s(x) \equiv B_s(w/2 - x, w) - e^{\pi i s}B_s(w/2 + x, w). \tag{9}$$

Using the fact that $B_s(x, w)$ is a solution of the difference equation (3), one arrives at the conclusion that

$$\phi_s(x + w) - \phi_s(x) = 0, \qquad (x \subset X). \tag{10}$$

When s is a positive integer or zero, it follows from well known properties of the classical Bernoulli polynomials that $\phi_s(x)$ is identically zero for all x. Thus, (i) $\phi_s(x)$ is a periodic function

* See Thiruvenkatacharya, *On some properties of the zeta function*, Journal of the Indian Mathematical Society, vol. 19 (1931), pp. 92–96.

† In §5 the existence of $B_s(x, w)$ and its analyticity with respect to s over the finite s plane was proved.

with period w for all values of s; (ii) when s is a positive integer or zero, $\phi_s(x)$ is zero for all values of x; (iii) $\phi_s(0)$ is given by

$$\phi_s(0) = (1 - e^{\pi i s})B_s(w/2, w) = (1 - e^{\pi i s})w^s B_s(1/2).$$

These properties point to a *significant role played by* $B_s(1/2)$ *in the behavior* of* $B_s(x, w)$.

8. *A Closely Related Integral.* Consider the integral† $L(x, s)$.

$$L(x, s) = \frac{w}{2\pi i}\int_{-i\infty}^{+i\infty}(x + z)^s\left(\frac{\pi}{w \cos \pi z/w}\right)^2 dz, \qquad (R(x) > 0). \tag{11}$$

We may write $L(x+w, s)$ in the form

$$L(x + w, s) = \frac{w}{2\pi i}\int_{w-i\infty}^{w+i\infty}(x + z)^s\left(\frac{\pi}{w \cos \pi z/w}\right)^2 dz. \tag{12}$$

The integrand of these integrals is analytic as a function of z over the region bounded by the two lines $R(z)=0$ and $R(z)=w$ except for a pole at $z=w/2$. The residue of this pole is $s(x+w/2)^{s-1}$. Hence the integral taken around the rectangle bounded by the two lines referred to above, and lines parallel to the axis of reals at $-iN$ and $+iN$ (N real and positive) comes to $ws(x+w/2)^{s-1}$. As $I(z)\to\pm\infty$, the absolute value of the integrand goes to zero like $t^s e^{-ct}$, $t\to\infty$, uniformly for $0\leqq R(z)\leqq w$ (c real and positive). It follows that if we let $N\to\infty$, the contributions to the circuit integral made over the above lines at $\pm iN$ approach zero. Thus, from (11) and (12), we have

$$L(x + w, s) - L(x, s) = ws(x + w/2)^{s-1}, \qquad (R(x) > 0). \tag{13}$$

Moreover, this relation is valid for all complex values of s. Also it is not difficult to show that for $R(s)<0$, $L(x, s)\to 0$ as $R(x)\to\infty$. Upon applying the reasoning used in proving Lemma 1, it follows that $L(x, s)$ is the unique solution of difference equation (13) for x on the right half-plane which satisfies the asymptotic condition (2). Hence, *for* x *on the right half-plane,* $L(x, s)$ *is identical with the generalized Bernoulli polynomial* $B_s(w/2+x, w)$.

* See N., p. 130.

† See N., p. 75.

9. *Relation to the Generalized Riemann Zeta-Function.* The generalized Riemann zeta-function may be introduced by the relation

$$\zeta(x, 1-s) = \sum_{k=0}^{\infty} \frac{1}{(x+k)^{1-s}}, \qquad (x \subset X,\ R(s) < 0). \tag{14}$$

For $R(x)>0$, the definition of $\zeta(x, 1-s)$ as a function of s has been extended to include the finite s plane and under such definition $s\zeta(x, 1-s)$ is found to be an entire function of s.*

On the other hand, from the difference equation satisfied by $B_s(x) \equiv B_s(x, 1)$, it follows that

$$\sum_{k=0}^{\infty} \frac{1}{(x+k)^{1-s}} = \frac{1}{s}\left[B_s(x+r+1) - B_s(x)\right], \qquad (x \subset X). \tag{15}$$

If $R(s)<0$, it is known that $B_s(x+r+1) \to 0$ as $r \to \infty$. Hence,

$$-\frac{1}{s} B_s(x) = \sum_{k=0}^{\infty} \frac{1}{(x+k)^{1-s}}, \qquad (x \subset X,\ R(s) < 0), \tag{16}$$

$$B_s(x) = -s\zeta(x, 1-s), \qquad (x \subset X,\ R(s) < 0). \tag{17}$$

When $R(x)>0$, the functions on each side of this relation, considered as functions of s, are defined and analytic over the finite s plane. Thus the above equation holds for all values of s, and we can state the following theorem.

THEOREM. *For $R(x)>0$ and for all values of s, the generalized Bernoulli polynomial of order one is related to the generalized Riemann zeta-function by equation* (17).

As far as the writer is aware, this is the first time that the relation (17) has been set up for values of s other than integers. The fact that it has been worth while to study the zeta-function as a function of a complex variable, may, by reason of equation (17), mean that the study of a generalized Bernoulli polynomial will yield further interesting and worth while results. In passing it is interesting to note that the relation (17) may serve to extend the domain of definition of $\zeta(x, s)$, for all values of s, to values of x over all of region X.

* See, for example, Whittaker and Watson, *Modern Analysis*, 3rd ed., Chap. 13.

From the relation between the zeta-function and the generalized Bernoulli polynomial there emerge several interesting consequences. From the well known functional relation for $\zeta(1, s)$ it follows that

$$sB_{1-s}(1) = 2^s\pi^{s-1} \sin(s\pi/2)\Gamma(2-s)B_s(1) \tag{18}$$

for all values of s. Since $B_s(1/2)$ plays a central role in the general behavior of the function $B_s(x, w)$, it is of interest to apply the above relation to $B_s(1/2)$. From the multiplication equation (6) when we set $w=1$, $n=2$, $x=1/2$, we obtain

$$B_s(1/2) = (2^{1-s}-1)B_s(1), \tag{19}$$

which, when substituted in (18), gives

$$B_{1-s}(1/2) = \frac{2^s\pi^{s-1}}{s}\left[\frac{2^s-1}{2^{1-s}-1}\right]\Gamma(2-s)\sin(s\pi/2)B_s(1/2). \tag{20}$$

This relation holds for all values of s.

If w were given other values than unity, a further extension of the zeta-function would result from the relation (17). This would mean that in the series definition of the zeta-function (14) the denominators $x+k$ would be replaced by numbers $x+kw$. Such a generalized zeta-function would be the solution of the difference equation

$$\frac{1}{w}[f(x+w, s) - f(x, s)] = -x^{-s}, \qquad (x \subset X), \tag{21}$$

analytic in s over the finite s plane (except for a pole at $s=1$) which for $R(s)>1$ has its asymptotic value equal to zero for $R(x)$ positively infinite.

ALBANY, N. Y.

THE METHOD OF MOMENT DISTRIBUTION FOR THE ANALYSIS OF CONTINUOUS STRUCTURES

BY LOUIS BRAND

1. *Introduction.* If ab is any member of constant cross-section forming part of a continuous structure, the moments at its ends, M_{ab} and M_{ba}, are given in terms of the angle-changes θ_a, θ_b at its ends and the lateral deflection per unit of length R by the slope-deflection equations*

$$M_{ab} = 2EK_{ab}(2\theta_a + \theta_b - 3R) \mp C_{ab},$$
$$M_{ba} = 2EK_{ba}(2\theta_b + \theta_a - 3R) \pm C_{ba}.$$

Here $K_{ab} = K_{ba}$ denotes the sectional moment of inertia of ab divided by its length (I/l) and C_{ab}, C_{ba} are the *numerical* values of the fixed-end moments due to the loading on ab. For the derivation of these equations and the sign conventions employed, reference may be made to the Bulletin just cited.

When there is no lateral deflection, or when this is neglected, $R=0$, and we write the slope-deflection equations

$$M_{ab} = 2EK_{ab}(2\theta_a + \theta_b) + M_{ab}^{(0)}, \tag{1}$$

$$M_{ba} = 2EK_{ba}(2\theta_b + \theta_a) + M_{ba}^{(0)}, \tag{2}$$

in which $M_{ab}^{(0)}$, $M_{ba}^{(0)}$ denote the fixed-end moments *inclusive of sign*. Suppose also that at all joints of the structure, other than certain fixed ends, there is no external momental load. At any such joint a, we must have

$$\sum M_{ai} = 0, \tag{3}$$

the summation ranging over all members that meet at a. At a fixed end c, θ_c is given. Owing to the continuity of the structure, all members meeting at a joint a rotate through the same angle θ_a. At each such joint we have an equation of type (3). Thus we have precisely as many equations (3) as we have unknown angles, so that in general these equations determine the angles

* Bulletin No. 108, Engineering Experiment Station, University of Illinois, p. 20.

uniquely. Having found them, equations (1) and (2) give the terminal moments for each member of the structure.

Hardy Cross has devised a method of successive approximations for solving equations (1), (2), and (3) for the terminal moments M_{ab} that is very simple and effective. In his original paper* he regards the method as the mathematical parallel of a physical process. To quote Cross: "The beams are loaded or otherwise distorted while the joints are held against rotation; one joint is then allowed to rotate with the accompanying distribution of the unbalanced moment at that joint while the resulting moments are carried over to the adjacent joints; then another joint is allowed to rotate while the others are held against rotation; and the process is repeated until the joints are 'eased down' into equilibrium."

2. *The Successive Approximations.* To begin the calculation corresponding to this process we write at the ends of each member the fixed-end moments corresponding to the load it carries. Consider a joint a at which the members ab, ac, $\cdots$, ak meet. If the algebraic sum of the fixed-end moments at a,

$$S = M_{ab}^{(0)} + M_{ac}^{(0)} + \cdots + M_{ak}^{(0)},$$

is not zero, we add a moment $-S$ at a and distribute it among the members meeting there in the proportion of their stiffness. The *stiffness* of any member ab is defined as the moment which must be applied at its end a to make $\theta_a = 1$ when b is held fixed ($\theta_b = 0$). From (1) it is clear that the stiffness of ab is $4EK_{ab}$. Thus the stiffnesses of the members ab, ac, $\cdots$, ak are proportional to K_{ab}, K_{ac}, $\cdots$, K_{ak}. The balancing moment $-S$ is therefore distributed among these members so that they receive respectively

$$-\frac{K_{ab}}{\sum K_{ai}} S, \quad -\frac{K_{ac}}{\sum K_{ai}} S, \cdots, \quad -\frac{K_{ak}}{\sum K_{ai}} S.$$

While joint a is allowed to rotate all other joints are held fixed: in particular, $\theta_b = \theta_c = \cdots = \theta_k = 0$. Hence from equations

* Proceedings of the American Society of Civil Engineers, May, 1930, pp. 919–928. For the paper with the complete discussions, see the Transactions of the American Society of Civil Engineers, vol. 96 (1932), pp. 1–156.

(1) and (2), a moment $4EK_{ab}\theta_a$ applied at end a of ab induces a moment $2EK_{ba}\theta_a$ at the fixed end b of ab. Thus, in the case of members of constant section, one-half of the moment distributed to ab in balancing the joint a is carried over to the end b. We therefore call 1/2 the *carry-over factor* for members of constant section.

The three basic operations involved in the above method of moment distribution are therefore:

BD (Balance and Distribute). *If the moments at a joint do not balance, add the balancing moment at the joint and distribute it to the members meeting there in the proportion of their stiffness.*

C (Carry-over). *One-half of each moment distributed to a member is carried over to its other end.*

The moments carried over in operation C destroy the balance achieved in operation BD. Hence BD must be repeated on the moments carried over. A succession of operations C and BD will be called a *cycle*. If, at the end of any cycle, the fixed-end moment at end a of member ab is added algebraically to all the distributed and carried-over moments there, the sum will give an approximate value of M_{ab}. We shall show that this approximate value approaches the true value of M_{ab} as the number of cycles performed increases indefinitely.

We next consider in detail this process of computing M_{ab}. The first cycle consists in writing the fixed-end moments at a and distributing the balancing moment at a according to operation BD. As regards the end a of ab, this gives the entries

$$M_{ab}^{(0)}$$

$$\text{(BD1)} \qquad M_{ab}^{(1)} = -\frac{K_{ab}}{\sum K_{ai}} \sum M_{ai}^{(0)}.$$

The moment $M_{ba}^{(1)}$, distributed to ab at b must now be carried over to a according to operation C. All these moments carried over to a create a new unbalance at this joint, which is then balanced and distributed according to operation BD. Thus the next two entries for the end a of ab are

$$\text{(C1)} \qquad \frac{1}{2} M_{ba}^{(1)},$$

(BD2) $$\frac{1}{2}M_{ab}^{(2)}, \quad \text{where} \quad M_{ab}^{(2)} = -\frac{K_{ab}}{\sum K_{ai}}\sum M_{ia}^{(1)}.$$

Similarly the next cycle of operations gives the entries at end a of ab:

(C2) $$\frac{1}{4}M_{ba}^{(2)},$$

(BD3) $$\frac{1}{4}M_{ab}^{(3)}, \quad \text{where} \quad M_{ab}^{(3)} = -\frac{K_{ab}}{\sum K_{ai}}\sum M_{ia}^{(2)}.$$

Continuing the cycles indefinitely and adding all the entries, we obtain

$$M_{ab} = M_{ab}^{(0)} + \left[M_{ab}^{(1)} + \frac{1}{2}M_{ab}^{(2)} + \frac{1}{4}M_{ab}^{(3)} + \cdots\right] + \frac{1}{2}\left[M_{ba}^{(1)} + \frac{1}{2}M_{ba}^{(2)} + \frac{1}{4}M_{ba}^{(3)} + \cdots\right]. \tag{4}$$

If the end a of ab is fixed, equation (3) does not apply and hence the operation BD is omitted. In this case M_{ab} is built up from $M_{ab}^{(0)}$ and the successive moments carried over from b, and the first series in (4) is absent.

3. *Convergence of Series.* We shall first establish the convergence of the series in (4). Let C denote the sum of the absolute values of all the fixed-end moments for the entire structure. Then writing $p_{ab} = K_{ab}/\sum K_{ai}$, we have from (BD1)

$$|M_{ab}^{(1)}| \leqq p_{ab}\sum |M_{ai}^{(0)}| < C.$$

Now at any joint a, other than a fixed end,

$$\sum |M_{ai}^{(1)}| = (p_{ab} + p_{ac} + \cdots + p_{ak})\left|\sum M_{ai}^{(0)}\right| \leqq \sum |M_{ai}^{(0)}|.$$

Forming such inequalities for all joints which are not fixed and adding the results we obtain

$$\sum |M_{ij}^{(1)}| \leqq \sum |M_{ij}^{(0)}| \leqq C.$$

Hence from (BD2)

$$|M_{ab}^{(2)}| \leqq p_{ab}\sum |M_{ia}^{(1)}| < C.$$

In the same way we may show that

$$| M_{ab}^{(n)} | < C, \qquad (n = 3, 4, \cdots).$$

Both series in (4) are therefore absolutely convergent.

4. *The Series Satisfy the Slope-Deflection Equations.* If we identify the first series in (4) with $4EK_{ab}\theta_a$, then the second series must represent $2EK_{ba}\theta_b$, and (4) reduces to equation (1). That this identification is justified is readily seen from the fact that after any operation BD, the summed moments at a joint a (which is not fixed) satisfy equation (3); for the operation BD was expressly designed to accomplish this. Hence in the limit the moments M_{ab} given by (4) satisfy the system of equations of type (3)—one equation for each free joint. Hence we see that the sums of the convergent series

$$M_{ij}^{(1)} + \frac{1}{2} M_{ij}^{(2)} + \frac{1}{4} M_{ij}^{(3)} + \cdots, \qquad (i \text{ not a fixed joint}),$$

satisfy precisely the same system of linear equations that $4EK_{ij}\theta_i$ satisfy. Thus if the system of equations (3) has a unique solution for the angles θ_i, this solution is given by

$$4EK_{ij}\theta_i = M_{ij}^{(1)} + \frac{1}{2} M_{ij}^{(2)} + \frac{1}{4} M_{ij}^{(3)} + \cdots. \tag{5}$$

5. *Moment at a Fixed End.* If a is the fixed end of a member ab, the equation (3) does not apply and the operation BD is therefore omitted. The fixed-end moment $M_{ab}^{(0)}$ and the moments carried over from b will total up to the actual external moment acting on the structure at a. For on putting $\theta_a = 0$ in (1) we have

$$M_{ab} = 2EK_{ab}\theta_b + M_{ab}^{(0)}; \tag{6}$$

and from (4), on omitting the first series due to the operation BD and retaining the second due to operation C,

$$M_{ab} = M_{ab}^{(0)} + \frac{1}{2}\left[M_{ba}^{(1)} + \frac{1}{2} M_{ba}^{(2)} + \frac{1}{4} M_{ba}^{(3)} + \cdots \right].$$

In view of (5) the series in brackets sums to $4EK_{ba}\theta_b$, so that

the limiting value of the right-hand side has precisely the value given in (6).

6. *Procedure for a Member Hinged at One End.* If a is the hinged end of a member ab, the end of a cycle will always reduce the total moment at a to zero. But this series of operations BD and C leading to the result zero may be avoided by taking the stiffness of ab proportional to $K'_{ab}=(3/4)K_{ab}$, and replacing the fixed-end moments $M^{(0)}_{ab}$, $M^{(0)}_{ba}$ at the outset by

$$H^{(0)}_{ab} = 0, \qquad H^{(0)}_{ba} = M^{(0)}_{ba} - \frac{1}{2} M^{(0)}_{ab}.$$

These are the terminal moments for ab due to the given loading when a is hinged and b fixed. No moments need then be carried over to a from b, and consequently there is nothing to carry back from a to b. To justify this procedure, we have from (1) and (2) in this case,

$$0 = 2EK_{ab}(2\theta_a + \theta_b) + M^{(0)}_{ab},$$
$$M_{ba} = 2EK_{ab}(2\theta_b + \theta_a) + M^{(0)}_{ba}.$$

On eliminating θ_a from these equations we obtain

$$M_{ba} = 2EK_{ab}\frac{3}{2}\theta_b + M^{(0)}_{ba} - \frac{1}{2} M^{(0)}_{ab}$$
$$= 4EK'_{ab}\theta_b + H^{(0)}_{ba}.$$

But from the successive approximations for obtaining M_{ba}, with $H^{(0)}_{ba}$ as a starting moment and no carry-overs from a, we have

$$M_{ba} = H^{(0)}_{ba} + \left[M^{(1)}_{ba} + \frac{1}{2} M^{(2)}_{ba} + \frac{1}{4} M^{(3)}_{ba} + \cdots \right],$$

where, in view of (5), the series in brackets is precisely $4EK'_{ba}\theta_b$ as all distributions to ba are now made proportional to K'_{ba}.

THE UNIVERSITY OF CINCINNATI

ON CERTAIN HIGHER CONGRUENCES*

BY LEONARD CARLITZ

1. *Introduction.* This note is concerned with the higher congruence

$$\prod_{\deg G<m} (t-G) \equiv A \pmod{P}. \tag{1}$$

Here A, P, G denote polynomials in an indeterminate x with coefficients in a Galois field $GF(p^n)$ of order p^n. The product in the left member extends over all G of degree less than some fixed m; the modulus P is assumed irreducible of degree k. As will appear below, we may without loss assume $k>m$.

The congruence (1) has either no solution at all, or else has p^{nm} distinct solutions; if t is any solution, then the general solution is furnished by $t+G$, where $\deg G<m$. Define σ_j by means of

$$(u+x)(u+x^{p^n})\cdots(u+x^{p^{n(m-1)}}) = \sigma_0 u^{m-1} + \sigma_1 u^{m-2} + \cdots + \sigma_{m-1}.$$

Put

$$P = x^k + c_1 x^{k-1} + \cdots + c_k,$$

$$P' = kx^{k-1} + (k-1)c_1 x^{k-2} + \cdots + c_{k-1},$$

$$F_{m-1} = (x^{p^{n(m-1)}} - x)(x^{p^{n(m-2)}} - x^{p^n})\cdots(x^{p^n} - x^{p^{n(m-2)}}).$$

Then we prove the criterion: *The congruence* (1) *is solvable if and only if each product* $(\sigma_j/(F_{m-1})^{p^n})AP'$, $(j=0,\cdots,m-1)$, *is congruent* (mod P) *to a polynomial of degree* $<k-1$.

2. *Some Properties of* $\psi_m(t)$. We denote by $\psi_m(t)$ the product appearing in the left member of (1). Also, we let

$$F_m = \prod_{i=0}^{m-1} (x^{p^{nm}} - x^{p^{ni}}), \quad L_m = \prod_{i=0}^{m-1} (x^{p^{n(m-i)}} - x), \quad F_0 = L_0 = 1.$$

Then†

* Presented to the Society, September 10, 1935.

† L. Carlitz, Duke Mathematical Journal, vol. 1 (1935), p. 141.

$$\psi_m(t) = \sum_{i=0}^{m} (-1)^{m-i} \begin{bmatrix} m \\ i \end{bmatrix} t^{p^{ni}}, \qquad \begin{bmatrix} m \\ i \end{bmatrix} = \frac{F_m}{F_i L_{m-i}^{p^{ni}}};$$

$$\psi_{m+1}(t) = \psi_m^{p^n}(t) - F_m^{p^n-1} \psi_m(t). \tag{2}$$

We shall require a generalized form of identity (2). From the definition of $\psi_m(t)$, it is clear that

$$\psi_{m+1}(t) = \prod_c (t + c_m x^m + c_{m-1} x^{m-1} + \cdots + c_0), \tag{3}$$

the product extending over all sets $(c_m, \cdots, c_0)$ in the $GF(p^n)$. The right member of (3) may be written in the form

$$\prod_{c_j} \prod_c{}' \{(t + c_j x^j) + (c_m x^m + \cdots + c_{j+1} x^{j+1} + c_{j-1} x^{j-1} + \cdots + c_0)\},$$

the inner product extending over $(c_m, \cdots, c_{j+1}, c_{j-1}, \cdots, c_0)$ only. If we put

$$f_j(t) = \prod_c{}' (t + c_m x^m + \cdots + c_{j+1} x^{j+1} + c_{j-1} x^{j-1} + \cdots + c_0),$$

then (3) becomes

$$\psi_{m+1}(t) = \prod_c f_j(t + c x^j). \tag{4}$$

Now by a formula due to E. H. Moore,* the polynomial $f_j(t)$ may be written as a quotient of two determinants:

$$f_j(t) = \frac{\Delta_j(t)}{D_j}, \tag{5}$$

where†

$$\Delta_j(t) = \left| t^{p^{n(m-i)}}, x^{m p^{n(m-i)}}, \cdots, x^{(j+1) p^{n(m-i)}}, x^{(j-1) p^{n(m-i)}}, \cdots, 1 \right| \tag{6}$$

where $i = 0, 1, \cdots, m$, and

$$D_j = \left| x^{m p^{n(m-1-i)}}, \cdots, x^{(j+1) p^{n(m-1-i)}}, x^{(j-1) p^{n(m-1-i)}}, \cdots, 1 \right|$$

* This Bulletin, vol. 2 (1896), p. 191.

† For brevity we write only the elements in the ith row: $|a_{i0}\, a_{i1} \cdots a_{is}|$, $(i = 0, 1, \cdots, s)$.

where $i=0, 1, \cdots, m-1$. By (5) and (6) it is clear that

$$f_j(t + cx^j) = f_j(t) + cf_j(x^j).$$

We shall evaluate $f_j(x^j)$. Using (6), we have

$$\Delta_j(x^j) = (-1)^{m-j} \left| x^{mp^{n(m-i)}}, \cdots, x^{jp^{n(m-i)}}, \cdots, 1 \right|, \qquad (i = 0, \cdots, m),$$

$$= (-1)^{m-j} F_1 F_2 \cdots F_m. \tag{7}$$

The calculation of D_j is somewhat more involved. Expanding the determinant

$$E_{m+1} = \left| t_{m-i}^{m-j} \right|, \qquad (i, j = 0, \cdots, m),$$

with respect to the elements of the first row, we have on the one hand

$$E_{m+1} = V_0 t_m^m - V_1 t_m^{m-1} + \cdots + (-1)^m V_m, \tag{8}$$

say. On the other hand, it is easy to see that

$$E_{m+1} = (t_m - t_0)(t_m - t_1) \cdots (t_m - t_{m-1}) E_m. \tag{9}$$

If now we put

$$t_m = t, \qquad t_i = x^{p^{ni}}, \qquad (i = 0, \cdots, m-1),$$

(8) and (9) evidently imply

$$D_j = V_j = V_0 \sigma_j^{(m)} = F_1 F_2 \cdots F_{m-1} \sigma_j^{(m)},$$

where $\sigma_j^{(m)}$ is the jth elementary symmetric function of the quantities $x^{p^{ni}}$, $(i=0, \cdots, m-1)$. Then by (5) and (7) we have the formula

$$f_j(x^j) = (-1)^{m-j} \frac{F_m}{\sigma_j^{(m)}}. \tag{10}$$

Returning now to (4), we see that

$$\psi_{m+1}(t) = \prod_c \{f_j(t) + cf_j(x^j)\},$$

and therefore, making use of (10), we have finally the generalized identity

$$\psi_{m+1}(t) = f_j^{p^n}(t) - \left(\frac{F_m}{\sigma_j^{(m)}}\right)^{p^n-1} f_j(t). \tag{11}$$

For $j=0$, this reduces to the identity (2).

3. *Necessity of the Criterion for Solvability.* Returning to the congruence (1), we see, by repeated applications of (2), that, for $m>k=\deg P$,

$$\psi_m(t) \equiv \psi_{m-1}^{p^n}(t) \equiv \psi_k^{p^{n(m-k)}} \pmod P.$$

In the case $m=k$, it is evident that

$$\psi_k(t) \equiv t^{p^{nk}} - t \pmod P;$$

and clearly the congruence

$$t^{p^{nk}} - t \equiv A \pmod P$$

is solvable only for $A=0$, in which case it has p^{nk} solutions. We may therefore assume without any loss in generality that $m<k$.

Making use of (11), we note that if (1) is solvable, then also the congruences

$$u_j^{p^n} - \left(\frac{F_{m-1}}{\sigma_j^{(m-1)}}\right)^{p^n-1} u_j \equiv A, \qquad (j = 0, \cdots, m-1), \tag{12}$$

are solvable. As for (12), we shall need the following theorem.

If we put

$$P = x^k + c_1x^{k-1} + \cdots + c_k,$$
$$P' = kx^{k-1} + (k-1)c_1x^{k-2} + \cdots + c_{k-1},$$

then the congruence

$$u^{p^n} - u \equiv A \pmod P$$

is solvable if and only if the product AP' is congruent (mod P) *to a polynomial of degree $<k-1$.**

Applying this criterion to (12) we have the following theorem.

THEOREM 1. *If* (1) *is solvable* ($m<k$), *then necessarily each product*

* Duke Mathematical Journal, vol. 1 (1935), p. 166.

$$\left(\frac{\sigma_j^{(m-1)}}{F_{m-1}}\right)^{p^n} AP', \quad (j = 0, \cdots, m-1), \tag{13}$$

is congruent (mod P) *to a polynomial of degree* $<k-1$.

4. *Sufficiency of the Criterion.* Making a slight change in notation, we rewrite (11) in the form

$$\psi_m(t) = f_j^{p^n}(t) - \alpha_j f_j(t), \qquad \alpha_j = \left(\frac{F_{m-1}}{\sigma_j^{(m-1)}}\right)^{p^n-1}; \tag{14}$$

we put

$$f_j(t) = \sum_{i=0}^{m-1} \beta_{ji} t^{p^{ni}}, \quad (j = 0, \cdots, m-1). \tag{15}$$

Let us now consider the system of congruences in $t_0, \cdots, t_{m-1}$:

$$\sum_{i=0}^{m-1} \beta_{ji} t_i \equiv u_j \pmod{P}. \tag{16}$$

The determinant $B = |\beta_{ji}|$ is easily calculated. By (14) we have the recursion

$$\beta_{j,i-1}^{p^n} - \alpha_j \beta_{ji} = \begin{bmatrix} m \\ i \end{bmatrix}, \qquad \beta_{j,m-1} = 1.$$

Then

$$\begin{aligned} B^{p^n} &= |\beta_{ji}|^{p^n} = |\beta_{ji}^{p^n}|, \qquad (i, j = 0, \cdots, m-1), \\ &= \left| \begin{bmatrix} m \\ 1 \end{bmatrix} + \alpha_j \beta_{j1}, \cdots, \begin{bmatrix} m \\ m-1 \end{bmatrix} + \alpha_j \beta_{j,m-1}, 1 \right| \\ &= |\alpha_j \beta_{j1}, \cdots, \alpha_j \beta_{j,m-1}, 1| \\ &= \frac{L_m}{F_m} |\alpha_j \beta_{j1}, \cdots, \alpha_j \beta_{j,m-1}, (-1)^{m-1} \alpha_j \beta_{j0}| \\ &= \frac{\alpha_0 \alpha_1 \cdots \alpha_{m-1}}{(F_1 F_2 \cdots F_{m-1})^{p^n-1}} B, \end{aligned}$$

and therefore

$$B = \frac{c}{\sigma_1 \sigma_2 \cdots \sigma_{m-1}} \frac{F_{m-1}^m}{F_1 F_2 \cdots F_{m-1}}, \qquad (c \text{ in } GF(p^n)). \tag{17}$$

In particular $B \not\equiv 0 \pmod{P}$, and hence for fixed u_j the system (16) has a unique set of solutions t_i. It remains to show that, if the u_j satisfy the congruence (12), then

$$t_{i+1} \equiv t_i^{p^n}, \qquad (i = 0, \cdots, m-2). \tag{18}$$

The proof follows closely the proof of (17). For simplicity we take only the case $i=0$ of (18); the general case may be treated in exactly the same way, but the notation is somewhat more complicated. We have

$$\begin{aligned} Bt_0 &\equiv \begin{vmatrix} u_j & \beta_{j1} & \cdots & \beta_{j,m-1} \end{vmatrix}, \\ Bt_1 &\equiv \begin{vmatrix} \beta_{j0} & u_j & \cdots & \beta_{j,m-1} \end{vmatrix}, \end{aligned} \qquad (j = 0, \cdots, m-1). \tag{19}$$

Then

$$\begin{aligned} B^{p^n} t_0^{p^n} &\equiv \begin{vmatrix} u_j^{p^n} & \beta_{j1}^{p^n} & \cdots & 1 \end{vmatrix} \\ &\equiv \left| A + \alpha_j u_j, \begin{bmatrix} m \\ 2 \end{bmatrix} + \alpha_j \beta_{j2}, \cdots, \begin{bmatrix} m \\ m-1 \end{bmatrix} + \alpha_j \beta_{j,m-1}, 1 \right| \\ &\equiv \begin{vmatrix} \alpha_j u_j, \alpha_j \beta_{j2}, \cdots, \alpha_j \beta_{j,m-1}, 1 \end{vmatrix} \\ &\equiv \frac{L_m}{F_m} \begin{vmatrix} \alpha_j \beta_{j0}, \alpha_j u_j, \alpha_j \beta_{j2}, \cdots, \alpha_j \beta_{j,m-1} \end{vmatrix} \\ &\equiv \frac{L_m}{F_m} \alpha_0 \alpha_1 \cdots \alpha_{m-1} B t_1. \end{aligned}$$

Comparing with (17) and (19), we see at once that $t_1 \equiv t_0^{p^n}$. We have therefore proved the following result.

THEOREM 2. *The criterion* (13) *furnishes a set of sufficient conditions that the congruence* (1) *be solvable. To carry out the solution we first solve the congruences* (12) *for* u_j, *substitute these values in the right member of* (16), *and solve for* $t_j \equiv t^{p^{nj}}$.

Note finally that the general solution of (1) is $t+G$, where t is a particular solution and G is any polynomial of degree $<m$; this follows from

$$\psi_m(t+G) = \psi_m(t) + \psi_m(G) = \psi_m(t).$$

5. *Another Criterion.* It is easily seen that the congruence

$$u^{p^n} - u \equiv A \pmod{P}$$

is solvable only when

$$A + A^{p^n} + \cdots + A^{p^{n(k-1)}} \equiv 0. \tag{20}$$

In any event the sum (20) is congruent to a quantity in $GF(p^n)$; we denote this residue by $\rho(A)$. Thus the congruence (12) is solvable only when

$$\rho\left\{\left(\frac{\sigma_j^{(m-1)}}{F_{m-1}}\right)^{p^n} A\right\} = 0, \qquad (j = 0, \cdots, m-1). \tag{21}$$

Consider now the sum

$$\begin{aligned}
\rho_m(M) &\equiv \sum_{j=0}^{m-1} (-1)^{m-1-j} x^{m-1-j} \rho(\sigma_j^{p^n} M) \\
&\equiv \sum_{j=0}^{m-1} (-1)^{m-1-j} x^{m-1-j} \sum_{i=0}^{k-1} \sigma_j^{p^{n(i+1)}} M^{p^{ni}} \\
&\equiv \sum_{i=0}^{k-1} M^{p^{ni}} \sum_{j=0}^{m-1} (-1)^{m-1-j} \sigma_j^{p^{n(i+1)}} x^{m-1-j} \\
&\equiv \sum_{i=0}^{k-1} M^{p^{ni}} (x^{p^{n(i+1)}} - x) \cdots (x^{p^{n(m+i-1)}} - x) \\
&\equiv \sum_{i=0}^{k-m} M^{p^{ni}} \frac{L_{m+i-1}}{L_i},
\end{aligned} \tag{22}$$

where for brevity we put

$$\sigma_j = \sigma_j^{(m-1)}, \qquad M = \frac{A}{F_{m-1}^{p^n}}.$$

Clearly (21) implies $\rho_m(M) = 0$. Conversely, it is easy to see that if $\rho_m = 0$, then (21) must hold. Indeed, since the sum (22) is congruent to a polynomial of degree $m-1$ (which is surely $< k$), its vanishing entails the vanishing of its m coefficients; these are precisely the left members of (21). Thus we have the following theorem.

THEOREM 3. *The congruence*

$$\psi_m(t) \equiv MF_{m-1}^{p^n} \pmod{P}$$

is solvable if and only if

$$\rho_m(M) \equiv \sum_{i=0}^{k-m} M^{p^{mi}} \frac{L_{m+i-1}}{L_i} \equiv 0.$$

It will be remarked that in general the first criterion is more useful. However, the latter is of some interest in itself. Furthermore, it suggests possible criteria for more general classes of congruences; I hope to develop the matter elsewhere.

DUKE UNIVERSITY

A DIFFERENTIAL EQUATION FOR APPELL POLYNOMIALS†

BY I. M. SHEFFER

By a *set* of polynomials $\{P_n(x)\}$, $(n=0, 1, 2, \cdots)$, we shall mean an infinite sequence in which $P_n(x)$ is of degree *exactly*‡ n. Corresponding to a given set $\{P_n\}$ there are infinitely many sequences of polynomials $\{L_n(x)\}$ (with $L_n(x)$ of degree not exceeding n) and sequences of numbers $\{\lambda_n\}$ such that $\{P_n\}$ satisfies the linear differential equation (usually of infinite order) with parameter:§

$$L[y(x)] \equiv \sum_{n=0}^{\infty} L_n(x) y^{(n)}(x) = \lambda y(x), \tag{1}$$

which for $\lambda=\lambda_n$ gives $P_n(x)$. In fact, suppose $\{P_n\}$ is given. Let $\{\lambda_n\}$ be any sequence of numbers subject only to the condition that λ_n is not identically zero in n. Then a unique sequence $\{L_n(x)\}$ exists such that $L[P_n(x)]=\lambda_n P_n(x)$, $(n=0, 1, \cdots)$, where not all the L_n's are identically zero, and where no $L_n(x)$ is of degree exceeding n. The polynomial $L_n(x)$ is readily obtained by recurrence from $L_0, \cdots, L_{n-1}$. If we write

$$L_n(x) = l_{n0} + l_{n1}x + \cdots + l_{nn}x^n, \tag{2}$$

---

† Presented to the Society, April 25, 1935.

‡ For many purposes it suffices to have $P_n(x)$ of degree not exceeding n. Here, however, it is convenient to use the stricter condition.

§ See Sheffer, American Journal of Mathematics, vol. 53 (1931), pp. 29–30, for a relation suggestive of (1).

where l_{nn} and other coefficients may be zero, we shall have

$$\lambda_n = l_{00} + nl_{11} + n(n-1)l_{22} + \cdots + n!l_{nn}. \tag{3}$$

If we *start* with equation (1), that is, with the coefficients $\{L_n(x)\}$, and ask for what values of λ there exists a polynomial solution, we find as a necessary condition that λ have one of the values (3). Conversely, if λ has any value λ_n of (3), then there will correspond a polynomial solution $P_n(x)$ which is unique to within a multiplicative constant, provided $\lambda_m \neq \lambda_n$, $(m \neq n)$. Moreover, $P_n(x)$ will be of degree exactly n, so that $\{P_n\}$ is a *set*. In the case in which $\lambda_m = \lambda_n$ for two (or more) distinct values of m and n, there may or may not be, for this common characteristic number, as many linearly independent polynomial solutions as the order of the characteristic number.

DEFINITION. Let us suppose that equation (1) is given. If for each λ_n given by (3) there is a polynomial solution $P_n(x)$ of degree (exactly) n, we shall say that equation (1) is *non-singular*.

COROLLARY. *Equation* (1) *is non-singular if* $\lambda_m \neq \lambda_n$, $(m \neq n)$.

If equation (1) is non-singular, but $\lambda_m = \lambda_n$ for some $m \neq n$, then if P_m and P_n are corresponding solutions of degree m and n, the polynomial $aP_m + bP_n$ is also a solution, for every choice of a and b. In this case the equation does not define a *unique*† set of polynomials $\{P_n\}$.

Because of the property of equation (1) that *any* set $\{P_n(x)\}$ is a set of solutions, (for a suitably chosen sequence $\{L_n(x)\}$), it seems appropriate to refer to (1) as a *universal type equation for sets of polynomials*. This comprehensive character of a type equation suggests that it may prove useful in the study of properties of sets of polynomials. We may propose the following program. Given a set of polynomials, or a class of sets; to characterize the type equations that they satisfy, and, in particular, to determine if (and when) they correspond to a type equation of *finite order*.

The present note investigates these questions for the class of Appell sets of polynomials.

The set $\{P_n\}$ is an Appell set if we have the relations

† We ignore the obvious arbitrary multiplicative constant attached to each $P_n(x)$.

$$\frac{dP_n(x)}{dx} = P_{n-1}(x), \qquad (n = 1, 2, \cdots). \tag{4}$$

The following are well known characterizations of Appell sets.

(i) *A necessary and sufficient condition that* $\{P_n\}$ *be an Appell set is that a formal power series*

$$A(t) \sim \sum_0^\infty a_n t^n \tag{5}$$

exist such that

$$e^{tz}A(t) \sim \sum_0^\infty P_n(x)t^n. \tag{6}$$

We shall refer to $A(t)$ as the *generating function* (or power series) for $\{P_n(x)\}$.

(ii) *Set* $\{P_n\}$ *is an Appell set if and only if there exists a sequence of numbers* $\{a_n\}$ *such that*†

$$P_n(x) = a_0\frac{x^n}{n!} + a_1\frac{x^{n-1}}{(n-1)!} + \cdots + a_n\frac{x^0}{0!}, \tag{7}$$

$$(n = 0, 1, \cdots).$$

THEOREM 1. *Let* $\{P_n\}$ *be any Appell set, with generating function* $A(t)$. *Define the (formal) power series*

$$B(t) \sim \sum_0^\infty b_n t^n \tag{8}$$

by

$$B(t) = \frac{A'(t)}{A(t)}. \tag{9}$$

Then $\{P_n\}$ *satisfies the equation of type* (1):

† In both (6) and (7), in order that $\{P_n\}$ be a true set, that is, for P_n to be of degree exactly n, it is necessary (and sufficient) that $a_0 \neq 0$. There is no great difficulty, however, if $a_0 = \cdots = a_k = 0$, $a_{k+1} \neq 0$. For then $P_n(x)$, as given by (6) and (7), is identically zero for $0 \leqq n \leqq k$, after which we get $P_n(x)$ of degree $n-(k+1)$; and if we define $B(t)$ by $A(t) = t^{k+1}B(t)$, then $b_0 \neq 0$, and the Appell set for $B(t)$ is the same set as for $A(t)$ after the first $(k+1)$ P's (which are identically zero) for $A(t)$ have been omitted.

$$L[P_n(x)] \equiv (b_0 + x)P_n'(x) + b_1P_n''(x) + b_2P_n'''(x) + \cdots = nP_n(x), \qquad (n = 0, 1, \cdots). \tag{10}$$

Conversely, if the set $\{P_n\}$ *satisfies the equation* (10), *where* $b_0, b_1, \cdots$ *are constants, then (to within multiplicative constants)* $\{P_n\}$ *is an Appell set, and its generating function* $A(t)$ *is defined by* (9).

From (3) we see that $\lambda_n = n$ for equation (10). Therefore $\lambda_m \neq \lambda_n$, $(m \neq n)$, so that to each λ_n corresponds an essentially unique polynomial, and it is of degree n. Suppose the set $\{P_n\}$ satisfies (10). Let $P_0(x) = a_0$, and choose the multiplicative constants so that in $P_n(x)$ the coefficient of x^n is $a_0/n!$. It is verified that $P_1'(x) = P_0(x)$. Assuming that $P_k'(x) = P_{k-1}(x)$ for $k = 1, 2, \cdots, n-1$, we shall prove it for $k = n$. Differentiating (10), we find that $L[P_n'(x)] = (n-1)P_n'(x)$. Hence, since L is non-singular, we have $P_n'(x) = c_nP_{n-1}(x)$. Comparing coefficients of x^{n-1}, we find that $c_n = 1$. Hence $\{P_n\}$ is an Appell set. Let its generating function be $A(t) \sim \sum_0^\infty a_nt^n$, where $a_0 = P_0$ as already stated. Then $P_n(x)$ has the representation (7). If we substitute this in (10) and equate coefficients of like powers of x, we get the equations

$$b_0a_{n-j-1} + b_1a_{n-j-2} + \cdots + b_{n-j-1}\ a_0 = (n - j)a_{n-j}, \qquad (j = 0, 1, \cdots, n).$$

But these relations are completely equivalent to relation (9), as is easily verified. The converse part of Theorem 1 is thus established.

Now let $\{P_n\}$ be any Appell set, and let $A(t)$ be its generating function. Let us define $B(t)$ by (9), so that the coefficients of (10) are defined. Let $\{Q_n\}$ be the Appell set defined by (10), of whose existence (and uniqueness if we choose $Q_0(x) = \alpha_0$) we have just given proof. Then the generating function $A^*(t)$ for $\{Q_n\}$ is defined, by the part already proved, by the equation $B(t) = A^{*\prime}(t)/A^*(t)$. Since $A(0) = A^*(0) = a_0$, therefore $A^*(t) = A(t)$. Hence $Q_n(x) \equiv P_n(x)$, $(n = 0, 1, \cdots)$, so that $\{P_n\}$ satisfies (10). This completes the proof.

COROLLARY. *A necessary and sufficient condition that a set* $\{P_n\}$ *be an Appell set is that it satisfy the recurrence relations*

$$(11)\quad nP_n(x) = (b_0 + x)P_{n-1}(x) + b_1P_{n-2}(x) + \cdots + b_{n-1}P_0(x), \qquad (n = 0, 1, \cdots).$$

For if $\{P_n\}$ is an Appell set, it satisfies (10) and therefore (11). Now suppose $\{P_n\}$ is a set satisfying (11). We verify at once that $P_1' = P_0$. Suppose $P_k' = P_{k-1}$, $(k = 1, 2, \cdots, n-1)$; to prove that $P_n' = P_{n-1}$. If we differentiate (11) and use our induction assumption, we get $nP_n' = nP_{n-1}$, so that $P_n' = P_{n-1}$, and the set $\{P_n\}$ is an Appell set.

Relations (10) and (11) give us two further characterizations of Appell sets. But they are not the only ones possible. For example, if we want an equation (for Appell sets) of type (1) with $\lambda_n = n(n-1)$, then as was proved in Theorem 1, we obtain the equation

$$(12)\quad L_{[2]}[P_n(x)] = n(n-1)P_n(x),$$

where the operator $L_{[2]}$ is defined by the equation

$$L_2[y] \equiv \sum L_{[2]n}(x)y^{(n)}(x)$$

with

$$(13)\quad L_{[2]}(x; t) \sim \sum_{n=0}^{\infty} L_{[2]n}(x)t^n = \frac{t^2}{A}[A'' + 2xA' + x^2A],$$

where $A(t)$ is the generating function for $\{P_n\}$.

More generally, if $\{P_n\}$ is an Appell set defined by $A(t)$, then

$$(14)\quad L_{[k]}[P_n(x)] = n(n-1)\cdots(n-k+1)P_n(x),$$

where

$$(15)\quad L_{[k]}(x; t) \sim \sum_{n=0}^{\infty} L_{[k]n}(x)t^n = \frac{t^k}{A}\left[A^{(k)} + \frac{k}{1!}xA^{(k-1)} + \frac{k(k-1)}{2!}x^2A^{(k-2)} + \cdots + \frac{k!}{k!}x^kA\right].$$

Now suppose that we wish λ_n to be a polynomial in n. We can suppose that† $\lambda_0 = 0$. Then we have

$$(16)\quad \lambda_n = nl_{11} + n(n-1)l_{22} + \cdots + n(n-1)\cdots(n-k+1)l_{kk},$$

† For this merely serves to alter the (constant) value of $L_0(x)$, and does not affect the polynomials $L_n(x)$ of index $n>0$.

where $l_{11}, \cdots, l_{kk}$ are constants (that is, independent of n). Exactly as in Theorem 1, we obtain the following result:

THEOREM 2. *Let $\{P_n\}$ be any Appell set and $A(t)$ its generating function. Let $\{\lambda_n\}$ be given by* (16) *for any $k \geqq 1$. Then $\{P_n\}$ satisfies the equation*

$$L[P_n(x)] \equiv \sum_{r=0}^{\infty} L_r(x) P_n^{(r)}(x) = \lambda_n P_n(x), \tag{17}$$

where L is defined by

$$\begin{aligned} L(x;t) &\sim \sum_{r=0}^{\infty} L_r(x) t^r \\ &= l_{11} L_{[1]}(x;t) + l_{22} L_{[2]}(x;t) + \cdots + l_{kk} L_{[k]}(x;t), \end{aligned} \tag{18}$$

the functions $L_{[p]}(x;t)$ being given by (15). *Conversely, given* (17) *with L defined by* (18) *for any function $A(t)$, then there exists an Appell set of polynomials† $\{P_n\}$ satisfying* (17), *and its generating function is precisely $A(t)$.*

We proceed to characterize the equations of type (1) that define one and the same Appell set.

THEOREM 3. *Let $\{P_n\}$ be an Appell set, and $A(t)$ its generating function. A necessary and sufficient condition that the set $\{P_n\}$ be a solution of an equation of type* (1) *(with $\lambda_0 = 0$, which is no restriction) is that the operator $L[y] \equiv \sum_{n=0}^{\infty} L_n(x) y^{(n)}(x)$ be defined‡ by*

$$L(x;t) \sim \sum_{n=0}^{\infty} L_n(x) t^n \sim \sum_{k=1}^{\infty} l_{kk} L_{[k]}(x;t), \tag{19}$$

where $L_{[k]}(x;t)$ is given by (15), *and the l_{kk}'s are constants. The characteristic numbers $\{\lambda_n\}$ are given by* (3).

† If $\lambda_m \neq \lambda_n$, $(m \neq n)$, then this Appell set gives *all* the polynomial solutions of (17) (to within multiplicative constants). But if m and n (distinct) exist such that $\lambda_m = \lambda_n$, this Appell set does not include all the polynomial solutions of (17).

‡ Although the last series in (19) may not converge, yet it serves to determine a unique sequence of polynomials $L_1(x), L_2(x), \cdots$ which are the coefficients of equation (1). For $L_{[k]}(x;t)$, expressed as a power series in t, begins with a term in t^k. Hence $L_n(x)$, which is the coefficient of t^n in $L(x;t)$, is given by a *finite sum*, namely, the coefficient of t^n in the sum $\sum_{k=1}^{n} l_{kk} L_{[k]}(x;t)$.

First we shall show that for any sequence of numbers $\{l_{kk}\}$ the Appell set $\{P_n\}$ satisfies equation (1) with L defined by (19). Let

$$L^{[n]}(x;t) = \sum_{k=0}^{n} l_{kk}L_{[k]}(x;t) \sim \sum_{r=0}^{\infty} L_r^{[n]}(x)t^r,$$

where n is any positive integer. The remark of the last footnote shows that $L_r^{[n]}(x) = L_r(x)$, $(r = 0, 1, \cdots, n)$. Now, by Theorem 2, we have

$$\begin{aligned} L^{[n]}[P_r(x)] &= \{rl_{11} + r(r-1)l_{22} + \cdots + r!l_{rr}\}P_r(x) \\ &= \lambda_r P_r(x), \qquad (r = 0, 1, \cdots). \end{aligned}$$

Hence

$$L[P_r(x)] = L^{[n]}[P_r(x)] = \lambda_r P_r(x), \qquad (r = 0, 1, \cdots, n);$$

and inasmuch as n can be chosen arbitrarily, it follows that $L[P_r(x)] = \lambda_r P_r(x)$ for *all* r. That is, $\{P_n\}$ satisfies equation (1).

Conversely, suppose the Appell set $\{P_n\}$ satisfies equation (1) (with $\lambda_0 = 0$); we shall show that L is defined by (19) for suitably chosen numbers $\{l_{kk}\}$. *Define* $\{l_{kk}\}$ by (3), and with this choice of $\{l_{kk}\}$, define an operator L^* by (19):

$$L^*(x;t) \sim \sum_{k=1}^{\infty} l_{kk}L_{[k]}(x;t),$$

where the function $A(t)$ of (15) is the generating function for $\{P_n\}$. From the half of the theorem already established, $\{P_n\}$ satisfies the equation $L^*[y] = \lambda^* y$, the set $\{\lambda_n^*\}$ being given by (3). But $\{\lambda_n\}$ also is given by (3), so that $\lambda_n^* = \lambda_n$. Subtracting the two equations, we find

$$(L - L^*)[P_n] = 0, \qquad (n = 0, 1, \cdots).$$

If we set $n = 0, 1, \cdots$ successively, we find that

$$L_0(x) - L_0^*(x) = L_1(x) - L_1^*(x) = \cdots = L_n(x) - L_n^*(x) = \cdots = 0.$$

That is, $L^* \equiv L$. This proves the converse.

We next determine the condition that an Appell set satisfy a finite order equation of type (1). Two well known examples are the polynomial sets $\{(x-a)^n/n!\}$ and $\{H_n(x)\}$, the latter being

the Hermite polynomials. They satisfy the following equations, respectively:

$$L[y(x)] \equiv (x - a)y'(x) = \lambda y(x), \qquad (\lambda_n = n);$$

$$L[y(x)] \equiv 2xy'(x) - y''(x) = \lambda y(x), \qquad (\lambda_n = 2n).$$

The corresponding generating functions are respectively $A(t) = e^{-at}$ and $A(t) = e^{-t^2/4}$.

THEOREM 4. *A necessary and sufficient condition that an Appell set* $\{P_n\}$, *with generating function* $A(t)$, *satisfy a finite order equation of type* (1) *is that* $A(t)$ *have the form*

$$A(t) = e^{Q(t)}, \tag{20}$$

where $Q(t)$ *is a polynomial.*

Suppose that $A(t)$ has the form (20), where Q is of degree k. Then from (9), $B(t) = Q'(t)$, a polynomial of degree $k-1$, so that $B(t) = b_0 + b_1 t + \cdots + b_{k-1}t^{k-1}$. The universal equation (10) for Appell polynomial sets therefore reduces to

$$\begin{aligned} L[P_n(x)] &= (b_0 + x)P_n'(x) + b_1 P_n''(x) + \cdots + b_{k-1}P_n^{(k)}(x) \\ &= nP_n(x), \qquad (n = 0, 1, \cdots), \end{aligned}$$

and this is an equation of finite order (namely, order k).

Conversely, suppose the Appell set $\{P_n\}$ satisfies a finite order equation of type (1); we shall show that its generating function $A(t)$ has the form (20). Let the equation be

$$L[P_n(x)] \equiv L_0(x)P_n(x) + \cdots + L_k(x)P_n^{(k)}(x) = \lambda_n P_n(x), \qquad (n = 0, 1, \cdots).$$

As already pointed out, we may suppose that $\lambda_0 = 0$ (so that $L_0(x) = 0$). If we write $L_n(x)$ in the form (2), then λ_n is given by (3), but since $L_n(x) \equiv 0$, $(n > k)$, λ_n reduces to the form (16). Let $L_{[k]}(x; t)$, $(k = 1, 2, \cdots)$, be the functions of (15), and define the operator L^* by $L^*(x; t) = \sum_{r=1}^{k} l_{rr} L_{[r]}(x; t)$, where $A(t)$ is the generating function for $\{P_n\}$. Then $\{P_n\}$ satisfies (by Theorem 2) the equation $L^*[P_n] = \lambda_n^* P_n$, where λ_n^* is determined by (16). Since (16) was used to determine λ_n, it follows

that $\lambda_n^* = \lambda_n$ for all n. Hence (as was argued before) we have $L^* \equiv L$. That is, the given equation for $\{P_n\}$ is defined by

$$L(x; t) = \sum_{r=1}^{k} l_{rr} L_{[r]}(x; t). \tag{a}$$

But to begin with we had

$$L(x; t) = \sum_{r=1}^{k} L_r(x) t^r. \tag{b}$$

Hence

$$\sum_{r=1}^{k} l_{rr} L_{[r]}(x; t) = \sum_{r=1}^{k} L_r(x) t^r. \tag{c}$$

Let

$$\frac{A^{(r)}(t)}{A(t)} = a_{r0} + a_{r1}\, t + \cdots, \qquad (r = 0, 1, \cdots, k). \tag{d}$$

Expanding the left member of (c) in a power series in t, and observing (from the right member) that powers of t higher than the kth degree cannot occur, we obtain the equations

$$\begin{aligned} & l_{11}(a_{1n}) + l_{22}(a_{2,n-1} + 2x a_{1,n-1}) + \cdots \\ & \quad + l_{kk}\left(a_{k,n-k+1} + \frac{k}{1!}\, x a_{k-1,n-k+1} + \cdots \right. \\ & \quad \left. + \frac{k(k-1)\cdots 2}{(k-1)!}\, x^{k-1} a_{1,n-k+1}\right) = 0, \quad (n \geqq k). \end{aligned} \tag{e}$$

These relations are to hold identically in x. Of the numbers $1, \cdots, k$, let s be the largest for which $l_{ss} \neq 0$. It then follows from (e) that

$$a_{1,n-s+1} = 0, \qquad (n \geqq k). \tag{f}$$

That is,

$$\frac{A'(t)}{A(t)} = B(t) = a_{10} + a_{11} t + \cdots + a_{1,k-s}\, t^{k-s},$$

so that (20) holds, where $Q(t)$ is a polynomial of degree $k-s+1$. This completes the proof.

If the Appell set $\{P_n\}$ satisfies an equation of type (1) of finite order, then there are infinitely many other equations (of type (1)) of finite order for which the set is a solution. For example, from $L[P_n]=\lambda_n P_n$ follows $L^2[P_n]=\lambda_n^2 P_n$, $L^3[P_n]=\lambda_n^3 P_n$, $\cdots$. This raises the question: What is the minimum order of all equations satisfied by $\{P_n\}$? The answer is given by the following statement.

THEOREM 5. *Let the Appell set $\{P_n\}$ satisfy a finite order equation of type* (1), *so that its generating function $A(t)$ is given by* (20). *Then the minimum order of all equations (of type* (1)) *satisfied by $\{P_n\}$ is precisely the degree† of the polynomial $Q(t)$.*

Let $Q(t)$ be of degree m. That the minimum order k cannot exceed m is shown by the proof of the first half of Theorem 4, so that $k \leqq m$. Let‡ $L[P_n]=\lambda_n P_n$ be the equation of minimum order k. From the proof of the second part of Theorem 4, we see that there is an integer s for which $1 \leqq s \leqq k$, and such that $Q(t)$ is of degree $k-s+1$. That is, $m=k-s+1$. But $k \leqq m$, so that $k \leqq k-s+1$. This requires that $s=1$, from which follows the desired conclusion that $k=m$.

PENNSYLVANIA STATE COLLEGE

† The single exception is when $Q(t)$ is of degree *zero*, in which case $A(t)$ is a constant and $P_n(x)=cx^n/n!$. This set satisfies a first order equation.

‡ As usual it is no restriction to suppose, as we do, that $\lambda_0=0$.

ON SOME EXTREMAL PROPERTIES OF TRIGONOMETRIC POLYNOMIALS WITH REAL ROOTS

BY J. GERONIMUS

1. *Introduction.* L. Fejér [1],† O. Szász [2], [3], [4] and E. v. Egerváry [5] have found many interesting extremal properties of non-negative trigonometric polynomials. In particular, Szász [3] has found that for every non-negative trigonometric polynomial of order $\leqq n$ with real coefficients,

$$(1)\quad G_n(\theta) = 1 + \Re\sum_{k=1}^{n} \bar{\gamma}_k e^{ik\theta}, \quad (\gamma_k = \alpha_k + i\beta_k;\ k = 1, 2, \cdots, n),$$

the inequality

$$(2)\qquad |\gamma_k| \leqq 2\cos\frac{\pi}{\left[\frac{n}{k}\right]+2}, \qquad (k = 1, 2, \cdots, n),$$

is valid.‡

The object of this note is to find the *minimum* of the modulus of the first coefficient γ_n, supposing that all roots of $G_n(\theta)$ are real. The first problem of this kind has been considered by Blumenthal [6]; we shall return in §4 to his problem and its generalization.

2. *Equality of Roots of $G_n^*(\theta)$ for Problem* 1. Consider the following problem.

PROBLEM 1. *Find the minimum of the modulus of the first coefficient γ_n of a non-negative trigonometric polynomial*

$$G_n(\theta) = 1 + \Re\sum_{k=1}^{n} \bar{\gamma}_k e^{ik\theta}$$

of order n with real roots.

---

† Numbers in brackets refer to the Bibliography at the end.

‡ $\Re z$ means real part of z; $[a]$ means the greatest integer $\leqq a$.

In order to solve this problem we shall prove the following simple lemma.

LEMMA 1. *All roots of the polynomial $G_n^*(\theta)$ for which the minimum in Problem* 1 *is attained must be equal.*

Consider a non-negative trigonometric polynomial†

$$(3)\quad G_n(\theta) = 4\sin^2\frac{\theta-\theta_1}{2}\sin^2\frac{\theta-\theta_2}{2}F_{n-2}(\theta) = \mathcal{R}\sum_{k=0}^{n}\bar{\gamma}_k e^{ik\theta},$$

where $F_{n-2}(\theta)$ is a non-negative trigonometric polynomial of order $n-2$ with real roots,

$$(4)\quad F_{n-2}(\theta) = \mathcal{R}\sum_{k=0}^{n-2}\bar{\gamma}_k^* e^{ik\theta} = \sum_{k=0}^{n-2}|\gamma_k^*|\cos(k\theta-\alpha_k),$$

where $\alpha_k = \arg\gamma_k^*$, $(k=0, 1, 2, \cdots, n-2)$, and $\alpha_0 = 0$.

On putting $\alpha=(\theta_1+\theta_2)/2$, $\delta=(\theta_1-\theta_2)/2$, we see easily that

$$(5)\quad \begin{aligned}\gamma_0 &= \gamma_0^*\left(1+\frac{1}{2}\cos 2\delta\right) - |\gamma_1^*|\cos(\alpha_1-\alpha)\cos\delta \\ &\quad + \frac{1}{4}|\gamma_2^*|\cos(\alpha_2-2\alpha); \qquad |\gamma_n| = \frac{1}{4}|\gamma_{n-2}^*|.\end{aligned}$$

We see that $|\gamma_n|$ does not depend on α, nor on δ; on the other hand γ_0 is maximal for $\delta=0$ if $\cos(\alpha_1-\alpha)\leqq 0$, or for $\delta=\pi$ if $\cos(\alpha_1-\alpha)\geqq 0$. In both cases the minimal value of $|\gamma_n|$ under condition $\gamma_0=1$ corresponds to $\delta=0$ or $\delta=\pi$; therefore θ_1 and θ_2 coincide.‡

3. *Polynomials for which γ_n has Extremal Values.* It follows from this lemma that $G_n^*(\theta)$ is

$$(6)\quad G_n^*(\theta) = C[1+\cos(\theta+\alpha)]^n,$$

α being an arbitrary real argument; it may be written thus:§

$$(7)\quad G_n^*(\theta) = \frac{C}{2^{n-1}}\left\{\frac{1}{2}C_{2n,n} + \sum_{k=1}^{n}C_{2n,n-k}\cos k(\theta+\alpha)\right\}.$$

† It is clear that all real roots of a non-negative trigonometric polynomial are of even multiplicity.

‡ θ_1 and $\theta_1+2\pi$ are not considered as different.

§ See [6], p. 392.

For this polynomial we have

$$\gamma_0 = \frac{1}{2^n} C(C_{2n,n}); \qquad |\gamma_n| = \frac{1}{2^{n-1}} C,$$

whence we find the ratio

$$\frac{|\gamma_n|}{\gamma_0} = \frac{2}{C_{2n,n}}. \tag{8}$$

We have proved the following theorem.

THEOREM 1. *If $G_n(\theta)$ is a non-negative trigonometric polynomial,*

$$G_n(\theta) = 1 + \Re \sum_{k=1}^{n} \tilde{\gamma}_k e^{ik\theta},$$

of order n with real roots, then

$$\frac{2}{C_{2n,n}} \leqq |\gamma_n| \leqq 1; \tag{9}$$

the maximum is attained for the polynomial†

$$G_{\max}(\theta) = 1 + \cos n(\theta + \alpha), \tag{10}$$

and the minimum for the polynomial

$$G_{\min}(\theta) = \frac{2^n}{C_{2n,n}} \{1 + \cos(\theta + \alpha)\}^n, \tag{11}$$

α being an arbitrary real argument.

4. *The Generalized Extremal Problem.* Consider now the following extremal problem.

PROBLEM 2. *Find the minimum of the ratio*

$$\frac{A_m^2 + B_m^2}{\lambda A_0^2 + \sum_{k=1}^{m} (A_k^2 + B_k^2)}, \tag{12}$$

† See [1], [2].

where

$$g_m(\theta) = A_0 + A_1 \cos\theta + B_1 \sin\theta + \cdots + A_m \cos m\theta + B_m \sin m\theta$$

is a trigonometric polynomial of order m with real roots, and λ *is an arbitrary non-negative number.*

The above mentioned problem of Blumenthal [6] corresponds to $\lambda = 1$. It is easy to see that for $\lambda = 2$ Problem 2 is a particular case of the Problem 1. Indeed we see that

$$g_m^2(\theta) = G_n(\theta) = \mathcal{R}\sum_{k=0}^{n} \tilde{\gamma}_k e^{ik\theta} \tag{13}$$

is a non-negative trigonometric polynomial of order $n = 2m$, while

$$\gamma_0 = A_0^2 + \frac{1}{2}\sum_{k=1}^{m}(A_k^2 + B_k^2), \qquad |\gamma_n| = \frac{A_m^2 + B_m^2}{2}; \tag{14}$$

therefore we have for $\lambda \neq 2$

$$\frac{A_m^2 + B_m^2}{2A_0^2 + \sum_{k=1}^{m}(A_k^2 + B_k^2)} \geqq \frac{2}{C_{2n,n}} = \frac{2}{C_{4m,2m}}. \tag{15}$$

To solve our problem for all $\lambda \geqq 0$ we shall put it in the following form.

PROBLEM 2′. *Find the maximum of the expression*

$$L(G) = \frac{1}{2\pi}\int_0^{2\pi} G_n(\theta)d\theta + \epsilon\left\{\frac{1}{2\pi}\int_0^{2\pi} [G_n(\theta)]^{1/2}d\theta\right\}^2, \tag{16}$$

$$(\epsilon \geqq -1),$$

where $G_n(\theta)$ *is a non-negative trigonometric polynomial*

$$G_n(\theta) = \mathcal{R}\sum_{k=0}^{n} \tilde{\gamma}_k e^{ik\theta}, \qquad (|\gamma_n| = 1),$$

of order $n = 2m$ *with real roots.*

5. *Equality of Roots of* $G_n^*(\theta)$ *for Problem* 2′. We shall prove the following lemma.

LEMMA 2. *All roots of the polynomial $G_n^*(\theta)$ for which the maximum in Problem 2' is attained must be equal.*

Put

$$[G_n(\theta)]^{1/2} = 2 \sin \frac{\theta - \theta_1}{2} \sin \frac{\theta - \theta_2}{2} F_{m-1}(\theta) \tag{17}$$
$$= [\cos \delta - \cos (\theta - \alpha)] F_{m-1}(\theta),$$

where $\alpha = (\theta_1 + \theta_2)/2$, $\delta = (\theta_1 - \theta_2)/2$, and $F_{m-1}(\theta)$ is a non-negative trigonometric polynomial of order $m-1$,

$$F_{m-1}(\theta) = \Re \sum_{k=0}^{m-1} \bar{c}_k e^{ik\theta}, \tag{18}$$

with real roots. Thus we get

$$\frac{1}{2\pi} \int_0^{2\pi} [G_n(\theta)]^{1/2} d\theta = c_0 \cos \delta - \frac{1}{2} |c_1| \cos (\beta_1 - \alpha), \tag{19}$$

where $\beta_k = \arg c_k$, $(k = 0, 1, \cdots, m-1)$, and $\beta_0 = 0$. Further let

$$F_{m-1}^2(\theta) = \Re \sum_{k=0}^{2m-2} \bar{c}_k^* e^{ik\theta}; \tag{20}$$

then we obtain

$$\frac{1}{2\pi} \int_0^{2\pi} G_n(\theta) d\theta = c_0^* \left(1 + \frac{1}{2} \cos 2\delta\right) \tag{21}$$
$$- |c_1^*| \cos (\beta_1^* - \alpha) \cos \delta + \frac{1}{4} |c_2^*| \cos (\beta_2^* - 2\alpha),$$

where $\beta_k^* = \arg c_k^*$, $(k = 0, 1, \cdots, 2m-2)$, and $\beta_0^* = 0$. Using (19) and (21) we have

$$L(G) = A \cos 2\delta + B \cos \delta + C, \tag{22}$$

where

$$A = \frac{1}{2} (c_0^* + \epsilon c_0^2),$$
$$B = - |c_1^*| \cos (\beta_1^* - \alpha) - \epsilon c_0 |c_1| \cos (\beta_1 - \alpha),$$
$$C = \frac{1}{4} |c_2^*| \cos (\beta_2^* - 2\alpha) \tag{23}$$
$$+ \frac{1}{4} \epsilon |c_1|^2 \cos^2 (\beta_1 - \alpha) + c_0^* + \frac{1}{2} \epsilon c_0^2.$$

It is important to point out that $|\gamma_n| = (1/4)|c^*_{2m-2}|$ does not depend on α, nor on δ. Since we have

$$c_0^* = c_0^2 + \frac{1}{2}\sum_{k=1}^{m-1} |c_k|^2, \tag{24}$$

it is clear that for $\epsilon \geqq -1$ we have

$$A = \frac{1}{2}(1+\epsilon)c_0^2 + \frac{1}{4}\sum_{k=1}^{m-1} |c_k|^2 > 0. \tag{25}$$

Therefore $L(G_n)$ is maximal for $\delta = 0$ if $B \geqq 0$, and for $\delta = \pi$ if $B \leqq 0$; in both cases θ_1 and θ_2 coincide, which proves our lemma.

6. *Polynomials having the Extremal Property.* We see that the polynomial $G_n^*(\theta)$ is

$$\begin{aligned} G_n^*(\theta) &= 2^{n-1}[1 + \cos(\theta + \alpha)]^n \\ &= \frac{1}{2}C_{2n,n} + \sum_{k=1}^{n} C_{2n,n-k} \cos k(\theta + \alpha), \end{aligned} \tag{26}$$

and we have for it

$$L(G_n^*) = \frac{1}{2}(C_{2n,n} + \epsilon(C_{n,n/2})^2). \tag{27}$$

Thus we have proved the following theorem.

THEOREM 2. *If $G_n(\theta)$ is a non-negative trigonometric polynomial of order $n = 2m$,*

$$G_n(\theta) = \mathcal{R}\sum_{k=0}^{n} \bar{\gamma}_k e^{ik\theta}, \qquad |\gamma_n| = 1,$$

with real roots, then

$$\begin{aligned} &\frac{1}{2\pi}\int_0^{2\pi} G_n(\theta)d\theta + \epsilon\left\{\frac{1}{2\pi}\int_0^{2\pi} [G_n(\theta)]^{1/2}d\theta\right\}^2 \\ &\qquad \leqq \frac{1}{2}(C_{2n,n} + \epsilon(C_{n,n/2})^2), \qquad (\epsilon \geqq -1); \end{aligned} \tag{28}$$

the maximum is attained for the polynomial

$$G_n^*(\theta) = 2^{n-1}\{1 + \cos(\theta + \alpha)\}^n, \tag{29}$$

α being an arbitrary real argument.

This result may also be stated as the following theorem.

THEOREM 2′. *If $g_m(\theta)$ is a trigonometric polynomial of order m,*

$$g_m(\theta) = A_0 + A_1\cos\theta + B_1\sin\theta + \cdots + A_m\cos m\theta + B_m\sin m\theta,$$

with real roots, then

$$\frac{A_m^2 + B_m^2}{\lambda A_0^2 + \sum_{k=1}^{m}(A_k^2 + B_k^2)} \geqq \frac{2}{C_{4m,2m} + \dfrac{\lambda - 2}{2}(C_{2m,m})^2}, \quad (\lambda \geqq 0); \tag{30}$$

this minimum is attained for the polynomial

$$g_m^*(\theta) = C\{1 + \cos(\theta + \alpha)\}^m. \tag{31}$$

BIBLIOGRAPHY

1. L. Fejér, *Über trigonometrische Polynome*, Journal für Mathematik, vol. 146 (1915), pp. 53–82.

2. O. Szász, *Über nichtnegative trigonometrische Polynome*, Sitzungsberichte der Bayerischen Akademie der Wissenschaften, (1917), pp. 307–320.

3. O. Szász, *Elementare Extremalprobleme über nichtnegative trigonometrische Polynome*, ibid., (1927), pp. 185–196.

4. O. Szász, *Über harmonische Funktionen und L-Formen*, Mathematische Zeitschrift, vol. 1 (1918), pp. 149–162.

5. E. v. Egerváry und O. Szász, *Einige Extremalprobleme im Bereiche der trigonometrischen Polynome*, ibid., vol. 27 (1927), pp. 641–652.

6. O. Blumenthal, *Einige Minimums-Sätze über trigonometrische und rationale Polynome*, Mathematische Annalen, vol. 77 (1916), pp. 390–403.

MATHEMATICAL INSTITUTE,
KHARKOW, U.S.S.R.

LINEAR CONNECTIONS OF NORMAL SPACE TO A VARIETY IN EUCLIDEAN SPACE*

BY C. B. TOMPKINS

1. *Introduction.* This paper deals with an extension of the Gauss formulas of a surface imbedded in ordinary space to apply to an m-dimensional variety imbedded in an n-dimensional euclidean space. The Gauss and Codazzi relations are extended to give integrability conditions in terms of the Christoffel symbols of the second kind of the variety, a set of tensors corresponding to the coefficients of the second fundamental form of a surface in three space and a set of non-covariant quantities: the connection coefficients of the normal space.

Several authors† have considered generalizations of the Frenet or the Gauss formulas to apply to varieties lying both in euclidean and in more general spaces, but the nature of some of the coefficients which may appear in these formulas seems to have escaped serious study.

In the considerations here there will be encountered an imbedding euclidean space of n dimensions, a variety of m dimensions lying in it and at each point of this variety an m-dimensional tangent space and an $(n-m)$-dimensional normal space. A vector in the imbedding space will be denoted simply by a letter, and all indices running from 1 to n will be suppressed. Latin indices lying between a and k will have the range from 1 to m and Latin letters from p through z when used as indices will have the range from 1 to $(n-m)$. Corresponding Greek letters will be used as summation indices.

The work in this paper differs from that in most of the previ-

* Presented to the Society, April 19, 1935.

† See, for instance, Voss, Mathematische Annalen, vol. 16 (1880), p. 129; H. Weyl, Mathematische Zeitschrift, vol. 12 (1922), p. 162; Schouten and van Kampen, Mathematische Annalen, vol. 105 (1931), p. 144; E. Bortolotti, Rendiconti del R. Istituto Lombardo di Scienze e Lettere, (2), vol. 64 (1931), p. 441; E. H. Cutler, Transactions of this Society, vol. 33 (1931), p. 832; C. E. Weatherburn, Reports of the Australian and New Zealand Association for the Advancement of Science, vol. 21 (1933), p. 12; Duschek and Mayer, *Lehrbuch der Differentialgeometrie*, vol. 2.

ously published papers in the method of selecting coordinates in the normal space; in this paper a set of perfectly arbitrary mutually orthogonal unitary vectors is set up as a coordinate system in the normal space at every point of the variety, while in most of the other papers the normal space is divided into several subspaces (the first containing all components of the first derivatives of the tangent vectors lying in the normal space, the second determined by additional components of the second derivatives, and so on) and in each of these subspaces an arbitrary coordinate system is introduced. Weyl derived the same Gauss formulas and integrability conditions we shall use, but he seems not to have noticed the properties of the coefficients which we shall call N_{pk}^{q}.

2. *The Gauss Formulas and Integrability Conditions.* A variety of m dimensions may be represented by a set of equations depending on m parameters, $x=x(u^k)$. In general there is a uniquely determined flat m-space tangent to this variety at every point; this is the linear space depending on the vectors $p_i \equiv \partial x/\partial u^i$; it is m-dimensional and uniquely determined if the vectors p_i are linearly independent. Also at every point there is a flat space which is absolutely perpendicular to the tangent space. If the tangent space is actually m-dimensional, the normal space is $(n-m)$-dimensional and contains sets of $(n-m)$ unit mutually orthogonal vectors. We may choose any such set and name it t_p. Now, any vector connected with a point of the variety may be written as a linear combination of the vectors p_i and t_p associated with that point. In particular we can write

$$\frac{\partial p_i}{\partial u^k} = \Gamma_{ik}^{\alpha} p_\alpha + L_{ik}^{\rho} t_\rho, \tag{1}$$

where the coefficients Γ_{ij}^{k} are the Christoffel symbols of the second kind connected with the variety and the choice of parameters.

We write also the equations for the partial derivatives of the vectors t_p,

$$\frac{\partial t_p}{\partial u^k} = -L_k^{\alpha p} p_\alpha + N_{pk}^{\rho} t_\rho, \tag{2}$$

where it is easily shown that $L^p_{ik} = g_{ia} L_k^{\cdot ap}$. Because the vectors t_p are mutually orthogonal unitary vectors, the coefficients N^q_{pk} must be anti-symmetric in the indices p and q.

The simplest method of arriving at integrability conditions seems to be that of variation of product integrals as developed by Rainich and Vaughan.* It is well known that we can take any set of equations of the form of (1) and (2) and by applying product integrals over some path arrive at a set of values for the vectors p_i and t_p along the path. However, for the variety to be a true variety it is necessary and sufficient that the vectors obtained for any point be the same no matter what path is used in deriving them; that is, the variation of the product integral used in their derivation must be zero. If we denote the matrices of the coefficients of the p_i and t_p in the expression for the partial derivatives with respect to u^k in (1) and (2) by A_k, a necessary and sufficient condition for the variation to vanish is

$$A_i A_k - A_k A_i + \frac{\partial A_i}{\partial u^k} - \frac{\partial A_k}{\partial u^i} = 0.$$

If we denote the matrices which make up A_k in the obvious manner by Γ_k, L_k, $L_k^{\cdot}$, and N_k, three relations analogous to the Gauss and Codazzi relations follow, conditions necessary and sufficient for the set of equations (1) and (2) to represent a true variety:

$$R^{\cdot}_{\cdot ki} \equiv \frac{\partial \Gamma_i}{\partial u^k} - \frac{\partial \Gamma_k}{\partial u^i} + \Gamma_i \Gamma_k - \Gamma_k \Gamma_i = L_k L_i^{\cdot} - L_i L_k^{\cdot},$$

$$(3) \qquad S^{\cdot}_{\cdot ki} \equiv \frac{\partial N_i}{\partial u^k} - \frac{\partial N_k}{\partial u^i} + N_i N_k - N_k N_i = L_k^{\cdot} L_i - L_i^{\cdot} L_k,$$

$$\frac{\partial L_i}{\partial u^k} - \frac{\partial L_k}{\partial u^i} + \Gamma_i L_k - \Gamma_k L_i + L_i N_k - L_k N_i = 0.$$

Having given the quantities Γ_k, L_k, and N_k satisfying these conditions and having given also a set of values for the components of the fundamental metric tensor at some point on the variety, the variety may be completely reconstructed except for position by means of a product integration followed by a Riemann integration.

* Abstract, this Bulletin, vol. 40 (1934), p. 233.

3. *Laws of Transformation.* All the coefficients of the equations (1) and (2) are invariant under transformations of the coordinate system of the imbedding euclidean space. Under transformations of the system of parameters u^k, the coefficients Γ_{ij}^k transform under the well known law for the transformation of connection coefficients; under these transformations, the coefficients L_{ik}^p transform as second rank doubly covariant tensors with indices i and k, while the similar quantities with one index raised must obviously transform as second rank mixed tensors. These transformations are all obvious from the form of equation (1). From equation (2) it is also evident that under arbitrary changes in parameterization of the variety the coefficients N_{pk}^q transform as first rank covariant tensors with index k.

In the normal space we allow rotations; more general transformations could be introduced with slight changes in the formulas already developed, but the introduction of such additional generality would simply necessitate the distinction between covariance and contravariance without rendering any useful service. We consider transformations of the type $\bar{t}=At$, where A is an orthogonal matrix. Applying this to the quantities of equation (1), we get the law of transformation for the coefficients L_{ik}^p,

$$\bar{L}_{ik} = AL_{ik}.$$

By continuing to equations (2), we find the law of transformation of the other coefficients,

$$\bar{N}_k = AN_kA^T - A\frac{\partial A^T}{\partial u^k}.$$

This law of transformation is formally exactly the same as the law which must be obeyed by coefficients of linear connection. This permits the statement of the following theorem concerning differentiation of complete tensors.*

THEOREM. *The operation of covariant differentiation on a complete tensor, in which operation the differentiation is with respect to one of the parameters u^k, and in which the quantities N_{pk}^q are used*

* That is, a set of quantities which are components of tensors in all the coordinate systems we are considering; the process we describe has been called complete differentiation (see E. H. Cutler, loc. cit.).

as connection coefficients for indices applying to the normal space, and in which the ordinary Christoffel symbols, Γ^k_{ij}, are used as connection coefficients for indices appertaining to the tangent space, yields a complete tensor.

The expression S^p_{qki} of the second of the integrability conditions (3) is simply the Riemann tensor of the connection coefficients of the normal space. The analogy between the first and the second of these conditions is striking. The third of these integrability conditions is rewritten under the definition of covariant differentiation just stated as

$$\Delta_k L^p_{ij} - \Delta_i L^p_{kj} = 0.$$

The anti-symmetry of the connection coefficients of the normal space reiterates the equivalence of covariance and contravariance.

4. *Vanishing of the Riemann Tensors.* It may be that there exists a choice of the vectors t_p such that the coefficients N^q_{pk} vanish simultaneously; from the law of transformation of these coefficients, it is evident that a necessary and sufficient condition for the existence of such coordinates is that the coefficients have the form $N_k = -A(\partial A^T/\partial u^k)$; or, since the matrix A is orthogonal, $\partial A/\partial u^k = N_k A$. This equation is of the type whose solution for the matrix A may be expressed by product integrals; we can find a solution over any particular path, but the solution must not depend on the path; that is, the variation of the product integral must vanish. A necessary and sufficient condition for this is the vanishing of the Riemann tensor of the normal connection coefficients, $S^p_{qki} = 0$. The solution we obtain by means of such an integration for A is an orthogonal matrix, for it can be shown that the product integral of any anti-symmetric matrix is orthogonal if the matrix used as an arbitrary constant of integration is orthogonal. Hence, *the vanishing of the Riemann tensor S^p_{qki} is a necessary and sufficient condition for the existence of coordinate systems for which all the N^q_{pk} vanish.*

The implications of the vanishing of this tensor are readily seen from the second of the integrability conditions (3); if we choose the parameters u^i in such a way that at a point in which

we are interested the p_i are unit mutually orthogonal vectors, then $L_k^{ip} = L_{ki}^p$, and the integrability condition becomes

$$\begin{vmatrix} L_{i\alpha}^{\gamma} & L_{k\beta}^{\gamma} \\ L_{i\beta}^{\alpha} & L_{k\alpha}^{\alpha} \end{vmatrix} = 0.$$

Now rotate the vectors so that one of the matrices, say $L_{..}^{\gamma}$, becomes a diagonal matrix, and the equation becomes

$$L_{ik}^{\alpha}(L_{ii}^{\gamma} - L_{kk}^{\gamma}) = 0.$$

This means that *the vanishing of the tensor S_{qki}^p implies that at every point on the variety there exists a set of mutually orthogonal vectors in the tangent space of such a nature that all the principal directions of all the Dupin indicating quadrics lie along these vectors*. This is evident when we notice that the equation implies either that the other matrices $L_{..}^{s}$ are all diagonal when $L_{..}^{\gamma}$ is or that $L_{ii}^{\gamma} = L_{kk}^{\gamma}$. If this last condition is true, the matrix $L_{..}^{\gamma}$ has no principal directions with respect to the indices i and k, and we may rotate the vectors p_i and p_k in their plane until the matrix $L_{..}^{s}$ becomes a diagonal matrix without destroying the diagonal property of $L_{..}^{\gamma}$.

In the case of a surface in a euclidean four-space we can get more definite results. The vanishing of the tensor S_{qki}^p implies that the indicatrix of Wilson and Moore* degenerates into a straight line. Indeed, if we choose the coordinate axis y along the vector t_1 and the axis z along t_2, the indicating conic may be written parametrically

$$y = L_{11}^1 \cos^2\theta + 2L_{12}^1 \sin\theta\cos\theta + L_{22}^1 \sin^2\theta,$$

$$z = L_{11}^2 \cos^2\theta + 2L_{12}^2 \sin\theta\cos\theta + L_{22}^2 \sin^2\theta.$$

Because of integrability conditions and the vanishing of the tensor S_{qki}^p, these reduce immediately to the equation of the line

$$\begin{vmatrix} x - L_{11}^1 & y - L_{11}^2 \\ L_{12}^1 & L_{12}^2 \end{vmatrix} = 0.$$

UNIVERSITY OF MICHIGAN

* Wilson and Moore, Proceedings of the American Academy of Arts and Sciences, vol. 52 (1916), p. 324.

OFFICIAL COMMUNICATIONS

Meetings of the Society have been fixed at the following times and places:

ST. LOUIS, MISSOURI ANNUAL MEETING, December 31, 1935–January 3, 1936.

This meeting will be held in conjunction with meetings of the A.A.A.S. and the Mathematical Association of America. Addresses will be delivered by Professor Szegö on *Some recent investigations concerning the sections of trigonometric and related series*, and by Professor Synge on *Tensorial methods in dynamics*. On Tuesday afternoon, at a joint session with the Mathematical Association and Section A of the A.A.A.S., Professor Carmichael will deliver his retiring address as Chairman of Section A, on *Linear differential equations of infinite order*. The Josiah Willard Gibbs Lecture will be delivered by Professor Vannevar Bush on the topic *Mechanical analysis*.

NEW YORK CITY, February 29, 1936.

Abstracts must be in the hands of Associate Secretary J. R. Kline, Low Library, Columbia University, New York City, not later than February 1. Abstracts received by the Secretaries on or before January 10 will appear in the January issue of this Bulletin. By invitation of the Program Committee, Professor E. Hopf will deliver an address entitled *On some metrically transitive systems*.

NEW YORK CITY, April 10–11, 1936.

Abstracts must be in the hands of Associate Secretary J. R. Kline, Low Library, Columbia University, New York City, not later than March 14.

CHICAGO, ILLINOIS, April 10–11, 1936.

Abstracts must be in the hands of Associate Secretary M. H. Ingraham, University of Wisconsin, Madison, Wisconsin, not later than March 14.

BERKELEY, CALIFORNIA, April 11, 1936.

Abstracts must be in the hands of Associate Secretary T. M. Putnam, University of California, Berkeley, California, not later than March 14.

CAMBRIDGE, MASSACHUSETTS, SUMMER MEETING AND COLLOQUIUM, August 31–September 5, 1936.

Abstracts must be in the hands of Secretary R. G. D. Richardson, Low Library, Columbia University, New York City, not later than July 27. The meeting will be held in connection with the celebration of the Harvard Tercentenary. A more detailed announcement will be printed in later issues.

R. G. D. RICHARDSON, *Secretary of the Society*.

Articles for insertion in the BULLETIN should be addressed to E. R. HEDRICK, Editor of the BULLETIN, University of California at Los Angeles. Reviews should be sent to W. R. LONGLEY, Yale University, New Haven, Conn. Notes should be sent to H. W. KUHN, Ohio State University, Columbus, Ohio.

Subscriptions to the BULLETIN, orders for back numbers, and inquiries in regard to non-delivery of current numbers should be addressed to the American Mathematical Society, 450–459 Ahnaip St., Menasha, Wis., or Low Memorial Library, Columbia University, New York City.

The initiation fees and the annual dues of members of the Society (see this BULLETIN, p. 322, May, 1930; and the List of Officers and Members, September, 1934, p. 66), are payable to the Treasurer of the Society, Professor G. W. Mullins, Low Memorial Library, Columbia University, New York City.

WHOLE NUMBER 435

CONTENTS

	PAGE
On the Complex Roots of Algebraic Equations. By A. J. KEMPNER	809
A Theorem on Higher Congruences. By LEONARD CARLITZ	844
The Hilbert Space Postulates—A Further Reduction. By A. E. TAYLOR	847
A Note on Lipschitz Classes. By E. S. QUADE	849
A New Proof of the Equivalence of E. H. Moore. By J. P. BALLANTINE	853
On the Bernoulli Distribution. By SOLOMON KULLBACK	857
On a Resultant connected with Fermat's Last Theorem. By EMMA LEHMER	864
Correspondences connected with a Pencil of n-ics. By A. R. WILLIAMS	868
Some Multiplication Theorems for the Nörlund Mean. By FLORENCE M. MEARS	875
On the Locus of an Analytic Equation in the Real Plane. By A. B. BROWN	881
A Partial Differential Equation connected with the Functions of the Parabolic Cylinder. By HARRY BATEMAN	884
A Generalization of the Bernoulli Polynomial of Order One. By B. F. KIMBALL	894
The Method of Moment Distribution for the Analysis of Continuous Structures. By LOUIS BRAND	901
On Certain Higher Congruences. By LEONARD CARLITZ	907
A Differential Equation for Appell Polynomials. By I. M. SHEFFER	914
On Some Extremal Properties of Trigonometric Polynomials with Real Roots. By J. GERONIMUS	924
Linear Connections of Normal Space to a Variety in Euclidean Space. By C. B. TOMPKINS	931

For official Communications and Notices, see the inside of the back cover.

GEORGE BANTA PUBLISHING COMPANY, MENASHA, WISCONSIN